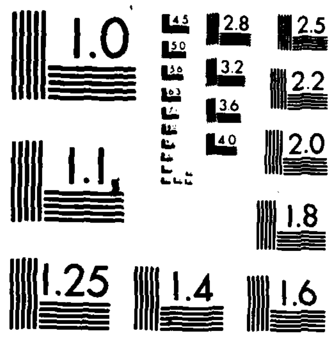
AD-A137 779 THE EVENT RELATED BRAIN POTENTIAL AS AN INDEX OF INFORMATION PROCESSING C..(U) ILLINOIS UNIV CHAMPAIGN COGNITIVE PSYCHOPHYSIOLOGY LAB E DONCHIN ET AL. OCT 83 1/2

UNCLASSIFIED CPL83-4 AFOSR-TR-84-0051 F49620-79-C-0233 F/G 5/10 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

COGNITIVE
PSYCHOPHYSIOLOGY
LABORATORY

Department of Psychology
University of Illinois
Champaign, Illinois 61820

Technical Report No. CPL83-4/
F49620-79-C-0233

October 1983

The Event Related Brain Potential as an Index of Information Processing, Cognitive Activity, and Skill Acquisition: A Program of Basic Research

FINAL Progress Report

Prepared for:
The Air Force Office of Scientific Research
Life Sciences Directorate

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER AFOSR-TR- 84-0051	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The event related brain potential as an index of information processing, cognitive activity, and skill acquisition: A program of basic research.		5. TYPE OF REPORT & PERIOD COVERED FINAL REPORT 01 SEP 79 - 31 AUG 83
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Emanuel Donchin, Christopher Wickens, and Michael G.H. Coles		8. CONTRACT OR GRANT NUMBER(s) F49620-79-C-0233
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Illinois, Department of Psychology Champaign, ILL 61820		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102 F 2313/A4
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research (NL) Bolling AFB, DC 20332		12. REPORT DATE OCT, 1983
		13. NUMBER OF PAGES 188
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Event-related brain potentials, P300, attention, workload, automaticity, skill, individual differences, Von Restorff effect, performance, reciprocity, reaction time, memory, error.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The materials assembled in this report represent work conducted with AFOSR support at the Cognitive Psychophysiology Laboratory during the reporting period. Appendix A of the report contains abstracts and papers that have been presented at various scientific meetings. In the text we present a brief review of these studies. For studies not included in Appendix A, a longer review is given. Appendix B gives a list of items that are either final versions of materials that were presented in previous progress reports or review chapters.

DD 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Table of Contents

INTRODUCTION

The materials assembled in this report represent work conducted with AFOSR support at the Cognitive Psychophysiology Laboratory (CPL) during the period 10/1/82-9/30/83. Appendix A of the report contains abstracts and papers that have been presented at meetings of the Society for Psychophysiological Research, the EEG and Psychophysiology Societies of Great Britain, the Evoked Potential International Congress (EPIC), and the Human Factors Society. In the text below, we present a brief review of these studies. For studies not included in Appendix A, a longer review is given. Appendix B gives a list of articles and chapters supported in whole or part by AFOSR. These items are either final versions of materials that were presented in previous progress reports or review chapters.

In the main, the CPL continued in this period to pursue closely related goals. The primary mission of our research is to develop an understanding of the Event Related Brain Potential (ERP) so that it can be used as a tool in the study of cognitive function and in the assessment of man-machine interactions. To this end, we are conducting studies that fall into four not altogether distinct categories, as follows:

A. The elucidation of the functional significance of the ERPs and application of this knowledge to an analysis of human cognitive function.

B. The use of ERPs in studies of workload.

C. The use of ERPs in the analysis of complex tasks and in the prediction of complex task performance.

D. Methodological studies.

Below, we present a systematic review of this research.

1. ERPs and Cognitive Function

In this section we focus on the elucidation of the functional significance of the ERPs and application of knowledge to the analysis of cognitive function. Much of this work focusses on the P300 component of the ERP. The noteworthy findings of the current period can be briefly summarized as follows:

1.1 P300 and Memory

We provided further support for the hypothesis that the P300 is a manifestation of those cognitive processes that affect representations in working memory. The current studies follow up on a study by Karis, Fabiani, and Donchin (in press, #A1) described in our last annual report. That study was designed to test the prediction that the probability that events will be recalled is proportional to the amplitude of the P300 they elicit. This prediction was confirmed by Karis et al. However, the results also emphasized that (a) many processes can be involved in memorization and recall only a subset of which may be related to the P300, and (b) there are considerable individual differences in the way subjects approach a recall-task, and these differences will strongly affect the relationship between P300 and recall.

Karis et al. used the von Restorff or Isolation paradigm (see #A1). The paradigm requires the subject to memorize a series of items. A deviant item is embedded in the middle of the series. It is commonly found that the deviant items (called "isolates") are recalled by the subject much better than comparable non-deviant items. This effect of enhanced recall of the

isolates is called **Von Restorff** or isolation effect. In our first experiment, we used series of words chosen at random, and isolated items by increasing, or decreasing, the size of the characters on the screen. As the isolates are both rare and task relevant they elicit large P300s. As we expected, in addition, the amplitude of the P300s varied across the isolates. We, therefore, could examine the relation between variance in P300 amplitude and the degree to which the isolates are recalled.

We ran 12 female subjects in this experiment. We presented lists of 15 words each. At the end of each list, the subject wrote down the words from that list that she could remember. No word was ever repeated. The ERPs elicited by the words were recorded.

Indices of the magnitude of the von Restorff effect and of the overall performance in the recall test were computed for each subject. We found striking individual differences in the degree to which subjects showed the von Restorff effect. We divided subjects into three groups, according to

RECALL + - + von R

the magnitude of their von Restorff effect. Subjects in one group showed enhanced recall for the isolates, but in general they showed very poor recall. They reported to have used rote strategies (that is, mere repetition of the words) to memorize the words. For these subjects, isolates that were recalled elicited larger amplitude P300s than isolates that were not recalled. On the other hand, subjects who did not show the von Restorff effect proved very good in recall. These subjects adopted an elaborative approach utilizing mnemonic devices to help their recall. In these subjects, there was no relation between P300 amplitude and later recall. Thus, the strategy of the subjects influenced both overall recall <u>and</u> the relationship between P300 amplitude and recall.

On the basis of these results, we proposed a 3-phase model. When an isolate is presented, a "memory updating" subroutine is invoked. We assume that P300 is an index of the activation of this subroutine. We also assume that all the subjects behave similarly in this phase. Strong individual differences show up in phases 2 and 3 of the model. For subjects who use rote strategies to memorize the words, the memory representations of the words are poorly organized and physical cues related to the isolation are very helpful for the retrieval of the words. For this reason, these subjects show a large von Restorff effect, and a strong relationship between P300 amplitude and memory. For subjects who use elaborative strategies to memorize the words, word representations in memory are very well organized and the physical cues provided by the isolation are useless for the word retrieval. Therefore, these subjects show little or no von Restorff effect but very high general performance, and the P300-memory relationship is obscured by the further processing.

1.1.1 Manipulation of Memorization Strategies

A straightforward test of this model consists of the manipulation of subjects' strategies to see if the pattern of results that we observed in different subjects can be reproduced within the same subject operating under different instructions.

This is what we are now doing. We use the von Restorff paradigm that was described before, but in two of the experimental sessions we give the subject explicit instructions about the strategies to use to memorize the words. Preliminary analysis of the data from a few subjects suggest the following conclusions. First, subjects will change their strategies

following instructions. Second, when they use the rote strategy, the von Restorff effect is large, overall recall is low, and P300 is related to later recall. Conversely, when the same subject uses elaborative strategies, the von Restorff effect is small, overall performance is high, and there is no relationship between P300 and later recall. These data are, of course, preliminary. However, they do suggest that P300 is related to a particular kind of memorial process.

1.1.2 Incidental Memory

There is another less direct way to test the model described above. The model suggests that, if subjects are not instructed to memorize materials, they will not engage in elaborate rehearsal. In this case, then, we can expect that the relationship between P300 and recall would hold for most people. This prediction was tested in study #A2 described in the appendix. Subjects were given an unexpected recall test after one of a series of oddball experiments. The stimuli were male and female names. One of the two categories was rare, with a probability of .20. No name was ever repeated. The subjects (n=35) were instructed to count either the rare or the frequent names. After this oddball was over, an unexpected free recall test was administered: the subjects were asked to write down as many names--both male and female--as they could remember. The data clearly show that names later recalled elicit a larger P300 than names later not recalled. This result holds for most of the subjects.

why is there variance in first block? Suggests to me intervening variable

In conclusion, these two experiments suggest that P300 amplitude to the stimulus when it is presented is related to subsequent memory performance, when further processes do not obscure this relationship. We interpret our

results as supportive of a model assuming that P300 amplitude is an index of memory updating.

1.1.3 Sternberg Experiment

Our second approach to the analysis of ERPs and memory has involved the use of the Sternberg paradigm.

Sternberg (1966) reported a study in which subjects memorized 1 to 6 digits and then were shown a probe digit and asked to report whether or not it was one of the digits memorized. He found reaction time (RT) increased linearly as a function of the number of digits in the memorized set. Using additive factors logic, Sternberg decomposed the RT into four processing stages: stimulus encoding, serial comparison, binary decision, and response execution. He interpreted the slope of the regression line to indicate the time necessary to make a memory comparison, i.e., the serial comparison stage. He interpreted the intercept of the regression line to reflect all other processes. Because the slopes for the positive and negative responses were identical, Sternberg argued that subjects perform an exhaustive search of the items in the memory set. This implies that subjects continue to scan the memory set even after a match has been detected.

In our experiment, we obtained ERP measures to the probe stimuli in order to try to understand the memory processes involved in the Sternberg paradigm. Forty five subjects were presented with memory sets ranging from one to five letters. Thirty probes were then presented, one every two seconds, and subjects were to determine if the probe matched one of the elements in the memory set. Subjects were instructed to respond by pressing one of two buttons as rapidly as possible without making errors.

Reaction time increased linearly as a function of set size for positive and negative probes. Negative probes were associated with longer reaction times than positive probes. The slope of the regression lines for positive and negative probes were essentially the same. The standard deviation of RTs increased as a function of set size for both positive and negative probes. Error rates for all conditions were under 5%. These results are consistent with the findings reported by Sternberg, i.e., an exhaustive search process.

Preliminary analysis of the ERP data has revealed that P300 amplitude is larger, and P300 latency is shorter, for positive than negative probes. P300 latency increases as a function of set size for positive but not negative probes. Note that this latter finding represents a clear dissociation between RT and P300 latency. For positive probes, both P300 latency and RT increase with set size. Furthermore, on a within-subject basis, P300 and RT are modestly but significantly correlated. For negative probes, however, RT, but not P300 latency, increases with set size. And, on a within-subjects basis, RT and P300 latency are not significantly related.

We interpret these data in the following way. The subject holds the items to be remembered in a "memory stack" - other letters of the alphabet may also reside in the stack but they are at a lower level than the positive items. For both positive and negative probes, the production of an RT response depends on a search through the positive items at the top of the stack. If a match between probe and item is made, the subject responds "yes" - if no match is made, the subject responds "no". In both cases, the subject apparently searches through all positive items before a response is

made. This accounts for the reaction time data. The P300, on the other hand, is dependent on a different process - namely, the matching of probe with an item in the stack. For positive probes, this match will be made faster and with more certainty, since the item is near the top of the stack. For negative probes, the match will be slower since the item to matched is lower down in the stack. Note that, in the case of the posi ve probes, the processes associated with RT and P300 are coupled - hence, /P300 latency correlation. For negative probes, however, RT and P300 are related to quite different processes - RT depends on the failure to find a match within the positive item set, while P300 depends on the presence of a match with an negative item. Hence, the decoupling of RT and P300.

This experiment is an excellent example of the merits of the psychophysiological approach in that measures of the ERP reveal more about cognitive processes than simple reaction time measures.

1.2 P300 and Error Detection

We have explored the functional significance of the ERP under circumstances in which the subject makes an error in responding to a stimulus (see Appendix #A3).

A tantalizing result that recurred in many of our studies in mental chronometry has been that on trials on which the subjects appear to be hasty in responding the P300 latency tends to be unusually long. This pattern appeared first in the study reported by Kutas, McCarthy, and Donchin (1977). The subjects were instructed to count the number of times names of males appeared in a list of common names. Some 80% of the names on the list were names usually ascribed to females. When the subjects were urged to be as

fast as possible they tended to press with a very short reaction time on the "female" button, even when the name presented was a "male" name. Strikingly, all these fast guesses were associated with long P300 latencies.

The conditions of the first study did not provide for the occurence of a large enough number of these trials to allow for very firm conclusions. McCarthy, Kutas and Donchin replicated the study using a much larger number of trials and urging the subjects even more to be fast. Indeed the number of errors increased greatly. The subjects appeared to be very biased to respond by pressing the female button. Again, the results suggested that for all subjects, the P300 latency was increased on these error trials (for details see McCarthy & Donchin, 1980). There remained, however, a number of questions. It was not possible, for example, to determine if the increased latency was due to the fact that an error was committed or to the fact that the response tended to be fast on these trials. It was also not possible to determine from these data the extent to which the emphasis on speed was critical for the pattern of results. Some investigators (e.g., Rossler, 1982) doubted that the component we identified was a delayed P300. It was suggested that the delayed peak represents a new component rather than a delayed P300.

We decided, therefore, to conduct a very detailed investigation that would try, in the design of the experiment, to address most of these concerns. To this effect we have run 7 male subjects, each in 4 conditions obtained by combining two levels of probability (p[male] = .50 and .20) and two instruction regimes (speed and accuracy). Data were recorded on 800 trials in each of the 4 cells from each of the four subjects. The ERPs were recorded, using Burden electrodes, from Fz, Cz, Pz, C1 and C2, all

electrodes referred to linked mastoids. Standard procedures were used to monitor EOG and EMG artifacts.

The data on the subjects' overt responses could be summarized as follows:

- The subject appeared to have adopted the instructional regimes as they tended to respond faster when instructed to be fast. Reaction times were longer, and the errors fewer when accuracy was emphasized.
- In the Speed conditions, the subjects hardly ever pressed the MALE button in response to a Female name. They made a substantial number of errors in response to Male names (i.e., they pressed the FEMALE button in response to Male names).
- The reaction times associated with these error trials were in general very fast. Correct responses to Male names were considerably slower.
- The reaction times to Female names were in general as fast as were the reaction times to Male names. Though, in both cases there was a distribution of reaction times.

It seems from the above, and from analyses that we do not have the space to describe in this report, that the subjects' behavior suggests that in both the Speed and the Accuracy conditions a bias to press the FEMALE button was maintained. Subjects' responses were thus driven largely by this bias. Alternate models were tested and were not consistent with all aspects of the data set.

The ERP data can be summarized as follows:

- The Male names in all series elicited a substantial P300, characterized by the scalp distribution commonly observed for the P300.
- Female names elicited a very small and indistinct P300 when the probability of such names was on .80.
- The latency of the P300 elicited by Male names was considerably longer when the subject erred on the trial than it was when the subject was correct. That is, for those male names that were responded to slowly, and correctly, the P300 latency was shorter than it was on those trials in which the subject responded very fast.
- Female names that were responded to with equal speed as were the error triggering male names did not elicit a delayed P300. In other words, it is unlikely that the longer P300 on error trials is due merely to the fast responses made on these trials.

The data described above lends support to a model that interprets the P300 as a manifestation of model revisions performed in Working Memory. According to this view the elicitation of the P300 is delayed on the error trials because the system is aware of the error and engages in additional processing before the trial information can be accomodated in the subject's world model.

1.3 Serial Stage Versus Continuous Flow Models

In appendix #A4, we describe an experiment that was designed in part to use psychophysiological measures to evaluate different models of human information processing. In this experiment, we used the measure of P300 latency to assess the time it takes a subject to evaluate a stimulus. We also used measures of the electromyogram and "sub-threshold" behavioral responses to define different types of trials in terms of the degree of error present. Specifically, in a choice reaction time task, we find that subjects sometimes initiate responses with the incorrect hand, although the complete response is actually made with the correct hand. These trials may be thought of as "partial" error trials. Subjects were required to make a discriminative response to the center letter in a five letter stimulus array. For some arrays, the noise letters surrounding the center letter were the same as the center letter; for other, incompatible arrays, the noise letters were those associated with the opposite response. We find that there are more error and partial error trials for incompatible arrays. These errors and partial errors lead to a delay in the production of the correct response. Our data also show that as P300 latency increases, the probability of error increases, and that for a given P300 latency, the probability of error is greatest if the subject responds quickly. If we assume that P300 latency is a measure of stimulus evaluation time, then these data (and other data - see appendix #A4) support the notion that information is passed from a stimulus evaluation system to a response activation system before the evaluation process is completed. In this sense, our data are more consistent with continuous flow models of information processing than with serial stage models.

1.4 Automaticity

Several investigators have argued for the existence of two qualititatively different forms of information processing: automatic and controlled. In this experiment (see Appendix #A5), we use measures of the ERP to evaluate these forms. Specifically, we demonstrate that P300 latency decreases as automatic processing is developed. This suggests that stimulus evaluation time decreases as automaticity progresses. Furthermore, an effect of probability on P300 amplitude, that is evident before automatic processing has developed, is absent after extensive training. This suggests that memory updating processes are attenuated under automatic processing.

2. The Use of Measures of ERPs in the Analysis of Workload

2.1 An Electrophysiological Analysis of Dual Task Integrality

The primary purpose of the present study is to investigate the phenomenon of dual-task integrality. This phenomenon occurs when two separate, but concurrently performed tasks can be processed within the same resource framework. In most dual-task cases, increasing the difficulty of one task is assumed to consume resources which normally would be employed in the processing of the other task. Thus the representation of the resources between the two tasks is presumed to be reciprocal in nature.

This assumption of resource reciprocity represents one of the primary tenets of the secondary task method of cognitive workload assessment. However, under conditions of dual-task integrality the secondary task increases processing demands within the domain of the primary task. Therefore in the case of dual-task intergrality resource reciprocity is not

obtained.

The overlap of relevant attributes between the two tasks is proposed to account for the integrated processing of the tasks. Two parameters which have been previously shown to influence the degree of integrality between two dimensions within a single task will be employed in the present dual-task context. These variables are the relationship between primary and secondary task stimulus objects (same or different) and the degree of correlation between the two tasks (zero or .8). Thus the present study represents an attempt to extrapolate findings concerning integrality between dimensions in a single task to integrality between two separate tasks.

The methodology employed to achieve this purpose involves the recording of event-related brain potentials (ERPs) to discrete changes in the primary and secondary tasks. The degree of integrality of the two tasks will be explored within the framework of the reciprocity of resources between the primary and secondary tasks. A demonstration that resource reciprocity between primary and secondary tasks does in fact exist requires the manipulation of primary task difficulty. This will be accomplished by varying the order of the control dynamics of the primary pursuit step tracking task (first, first/second and second order dynamics).

In addition to investigating the two variables which may influence the degree of integrality between the concurrently performed tasks, the present study will also address the integrality issue within multiple resource theory by manipulating the resources presumably required for primary and secondary task performance.

In one case both the primary and secondary tasks will require substantial spatial processing while in the other condition the primary task will necessitate spatial processing while the secondary task will require that the subjects attend to the intensity of the relevant stimuli. In all conditions the primary task is a single axis pursuit step tracking task in which the subject is required to cancel the error between the target and cursor via the manipulation of a joystick. The secondary task involves covertly counting one of two events presented in a .5/.5 Bernoulli series.

Data are currently being analyzed. It is anticipated that the study will be completed by the end of January 1984.

2.2 P300 and Resource Reciprocity

This study was designed to explore further the utility of ERP components as indices of mental workload. Previous studies conducted in this laboratory have indicated that the P300 componant is sensitive to certain aspects of cognitive functioning related to workload. The majority of these studies have employed a dual task paradigm.

The assumption underlying this research is that a human operator has pools of resources at his disposal during the performance of a task. More difficult tasks are assumed to require more resources. Thus, the workload associated with a primary task is assessed in terms of measures, either behavioral or psychophysiological, associated with the secondary task. In other words, a difficult primary task will drain away resources that could otherwise be utilized by the secondary task.

This laboratory has concentrated on secondary tasks employing the "oddball" paradigm and measures of the ERP. Given that the amplitude of

P300 is proportional to the extent to which a subject allocates resources to the processing of a stimulus, it seems reasonable to suppose that the P300 component may serve as an index of the relative relevance of the oddball task. Thus, reductions in P300's generated by the secondary task tones that are related to increased primary task difficulty are presumed to reflect increased resource allocation to the primary task. Thus, because the amplitude of secondary task P300's declined as the number of elements to be monitored increased in a study by Heffley, he argued that P300 was sensitive to the increased perceptual workload of the primary task. Conversely, in a study by Isreal no decrements in secondary P300 amplitude were observed as the number of dimensions was increased from one to two within the context of a primary tracking task. This pattern of results has been interpreted as indicating that P300 is sensitive to increments in primary task difficulty when the difficulty manipulation lies within the perceptual domain, but not when the difficulty is manipulated within the sensori-motor domain.

Following the above logic, Kramer has interpreted similar dual task data as confirming the hypothesis that as the system order control is increased during a step tracking task (ie. from a velocity to an acceleration system) the demands on perceptual resources are increased. The locus of this effect is described as perceptual rather than sensori-motor because this increase in primary task difficulty was reflected in a reduction in secondary task P300 amplitude. The term "resource reciprocity" was coined to describe the situation in which decreases in secondary task P300's are associated with increases in P300's generated by the primary task. Such resource reciprocity was subsequently demonstrated by Kramer with regard to a step tracking task.

The present study was designed to determine whether the dissociations in P300 sensitivity outlined above could be replicated in the situation where dimensionality and system order were orthogonally manipulated. A step tracking task was developed in which subjects were run through four conditions (2 system orders x 2 dimensions) within the context of both single and dual task instructions (ie. tones either present or absent). Thus, there were a total of 8 different conditions for each subject. In addition, subjects also performed the oddball task in the absence of the primary tracking task. ERP measures were obtained for both primary task stimuli (a step change) and secondary task stimuli (a tone).

Preliminary data analysis has confirmed that both of the manipulations affected subject performance (in terms of RMS error) with an interaction between the dimension and order manipulations being present in the single task conditions. Performance was worst in the two dimensional, second order system.

The ERPs obtained from 26 subjects during performance of these tasks has been analyzed. The parietally maximal positive component approximately 500 msec following stimulus onset has been identified as the P300 component. Resource reciprocity has been observed for this component with respect to the order manipulation but not with respect to the manipulation of dimensionality. In other words, this component is larger for primary task step changes when the subject is operating under a second order control system and is smaller following the secondary task tone when the secondary task is performed in conjunction with a second order primary task. Such reciprocity was not observed for the dimensionality manipulation. Even

though there is a decrement in the P300's to the secondary task as a function of increasing the number of dimensions in the primary task, these secondary task P300 decrements are not accompanied by corresponding increases in primary task P300's. A more detailed analysis of these data awaits the completion of the collection of the entire data set.

3. Complex Tasks

In this section, we review four projects which, though not directly supported by this AFOSR contract are related to the aims of the AFOSR. These three projects all involve the use of a complex task, "Space Fortress", which was adapted by us from a video game. Briefly, the subject must maneuver a space ship, identify and evade mines, and fire lasers to destroy the mines and, ultimately, a space fortress.

DARPA?

3.1 Additive Factors and Task Analysis

The first project was designed to evaluate the use of the additive factors procedure for the analysis of complex tasks into their components. The procedure and the results of this project are described in detail in Appendix #A7. Briefly, we took subjects who had been given extensive training on the task (experts) and required them to perform the task under varying degrees of difficulty. Difficulty was manipulated by varying (a) the speed of the hostile elements, (b) the memory requirements associated with the correct identification of the hostile elements, (c) the difficulty of a motor response that was necessary to perform this identification, and (d) whether or not the hostile elements disappeared briefy from view. We measured 29 different aspects of the subject's performance and looked at the pattern of main effects and interactions relating the the various

manipulations of difficulty to the different performance measures. In particular, we looked for clusters of performance measures that showed a similar responsiveness to the particular experimental variables. The results revealed three major clusters of performance measures. These seem to be associated with appraisal processes, motor processes, and perceptual-motor processes. Thus, we argued that successful performance of the task requires at least three different skills, one associated with each process.

3.2 Long Duration Missions

The next project was designed to determine whether performance decrements due to continuous performance of the task for 12 hour "missions" would be (a) equivalent for all skills, (b) related to changes in various ERP components, (c) different for novice and expert subjects, (d) different for day and night missions. The results are given in detail in Appendix #A8.

3.3 Learning Strategies

The next project is attempting to determine whether the task analysis described under 3.1 above can be used to guide the selection of training regimes for acquisition of the complex task. Given that we identified three skills associated with performance of the task, we proposed that each of these skills can be acquired through "part" training. However, because the perceptual-motor process interacted with the other processes, we proposed that the skill associated with this process has to be acquired in a "whole" training regime. We also argued that this "whole" training regime should be adaptive - that is, difficulty should be gradually increased for

that aspect of the task (speed of the hostile elements) that is related to the perceptual-motor process. These predictions are currently being tested.

3.4 Prediction of Performance

Finally, we are determining the value of ERP measures as predictors of performance on both the complex task and various sub-tasks. The development of these subtasks has been aided by the additive factors analysis on the complete Space Fortress task (3.1).

We have developed an ERP battery that includes several well studied paradigms. From these we will obtain information on a variety of ERP components, including P300, the contingent negative variation (CNV), slow wave, and N200. For some components we will have information from several experimental paradigms. From these data we will try to develop a composite ERP score to predict overall Space Fortress performance. We will also examine the relationship between the individual ERP components in each paradigm and performance on the Space Fortress subtasks.

Almost 40 subjects have now completed four ERP sessions. The experiments in these sessions constitute our ERP battery. Half the subjects will repeat session 1 two additional times (after 1 week and 3 months), permitting us to address important questions on the reliability of ERP components. We have also included a separate session to administer a psychometric battery. The four ERP sessions include oddball paradigms, a CNV paradigm, a Sternberg paradigm, and a dual-task tracking paradigm.

3.4.1 Oddballs

A series of events that can be divided into discrete classes is called an "oddball" when one event (or class of events) is much rarer than the other (although sometimes even series with 50-50 probability are called oddballs). Since such paradigms have been used extensively, typical ERPs can be easily recognized, and subjects producing anomalous waveforms identified.

We are using both simple and complex visual oddballs, and an auditory oddball with choice reaction time. From these we obtain measures of both P300 amplitude and latency. We also test subjects' memory for the names used in the complex oddball, using both recall and recognition.

3.4.2 CNV Paradigm

In the CNV paradigm a letter (H or S) is sometimes presented at S1, and sometimes at S2. This letter indicates the hand to be used for responding (by squeezing a dynamometer). When the letter is presented at S1 the subject is able to prepare the response, executing it as soon as S2 appears, while when the letter appears at S2 the subject can only prepare for the perceptual decision and response that will be required at S2.

The CNV is sensitive to these preparatory states, and by including a go/no-go condition we will be able to partial out the response related components.

3.4.3 Sternberg Paradigm

The Sternberg paradigm has been described in detail in Section 1.1.3. In it we will focus on stimulus categorization and evaluation as indexed by P300 latency.

3.4.4 Step-Tracking

In previous step-tracking experiments performed in our laboratory subjects have counted rare tones while engaging in a one dimensional pursuit tracking task. P300 amplitude to the secondary task tones is sensitive to the perceptual and cognitive demands of the primary task. When tracking difficulty is increased by changing to a second order system, P300 amplitude to the tones decrease.

We are examining individual differences in this decrement in P300 amplitude, and are also including conditions that increase response load by requiring tracking in two dimensions. P300 amplitude to a secondary task tone does not change with increases in the dimensionality of the primary tracking task, but these two effects (dimensionality, system order) have never been observed in the same experiment.

From these paradigms, we will be able to extract several measures of P300 latency and amplitude. P300 latency is related to stimulus evaluation, while amplitude, in our paradigms, will reflect short term memory (via sequential effects), "context updating" (by examining the P300-memory relationship), and resource allocation and capacity (in step tracking, which involves dual task methodology). Information on preparation for perceptual processing and motor responses will come from the CNV paradigm.

The Space Fortress subtasks have been developed to require processing reflected by these ERP components. These ERP experiments will also provide information on concurrent information processing (step tracking), perceptual speed (Sternberg, auditory oddball), and time estimation and anticipatory behavior (CNV paradigms).

Throughout the project we will be collecting subjective estimates of workload and task demands. The utility of such measures is controversial, in part because of unreliable and unsophisticated methodology. Dr. Danny Gopher, an expert in this area, is collaborating with us on this aspect of the project.

We have administered the symbol digit modalities test (SDMT) at the start of each session to serve as a reference point for these subjective ratings. After each task subjects gave an estimate of the demands and workload imposed by the task. They did this by assigning the task a number after comparing it to the SDMT, which was assigned a value of 100.

The central question of this project concerns the predictive value of information derived from ERP measures for performance in the complex task, "Space Fortress". We approach this question in two ways. First, we have devised several subtasks, based on the additive factors analysis of the whole task (see above). We expect particular ERP measures to be related to performance on particular sub-tasks depending on the degree to which those skills required to perform the sub-task are related to the processes manifested by the ERP measures. Second, using a multiple regression analysis, we will determine the relative value of different ERP measures as predictors of performance on the whole task.

3.4.5 Space Fortress Subtasks

3.4.5.1 Aiming

The ship is in the middle of the screen, and a mine appears in one of 24 positions on the periphery. The subject must rotate the ship and fire at the mine.

3.4.5.2 Time Estimation

There is no ship on the screen. An "X" appears and the subject is instructed to make a double press as close to 225 msec as possible (although anything between 150 and 300 msec is acceptable). The actual time is displayed at the bottom of the screen after each double press.

3.4.5.3 Sternberg Task

Again, no ship is on the screen. There are four letters in the positive set, four in the negative. A letter appears. If it is a member of the positive set the subject must execute a double press at the appropriate rate (150-300 msec), and then fire a missile (no aiming is required, as nothing is on the screen). For other letters no double press is required, and the subject must only fire a missile.

3.4.5.4 Flying the Ship

The ship starts to move and the subjects task is to stop it. This requires rotating the ship until it points in a direction opposite to its motion, and applying thrust sufficient to stop it, but not enough to accelerate it in the new direction.

4. Technical and Methodological Advances

We have continued to pursue our interest in methodological and technical advances which aid in the quantfication and analysis of ERPs.

4.1 Vector Filters

Following our solution of the eye-movement artifact problem (see last year's report), we have turned our attention to another problem in the analysis of ERPs. This problem concerns the quantification of a component of the ERP when the definition of the component includes a distributional aspect. For example, the P300 is defined, not only in terms of its polarity and latency, but also in terms of its distribution across different scalp locations. It is seen most positively at the parietal electrode and least positively at the frontal electrode. The critical question is - how do you quantify distributional information ?

In Appendix A6, we describe a method which permits the assessment of the degree of similarity between an obtained ERP distribution and a distribution defined, a priori. Thus, for a single ERP trial, or for an average ERP, we can measure the "P300ness" of each point in the waveform. This procedure can be conceptualized as filtering the ERP for its distributional characteristics. This "vector filter" procedure permits an asesseement of both P300 amplitude (the maximum value of the filter output) and latency (the timepoint of this maximum) for both single trials and average ERPs.

4.2 The Consistency of ERPs

We have begun to evaluate the consistency of various aspects of the ERP since a basic issue in the application of ERPs in the assessment of human operators is the consistency across situations of the ERP generated by a given operator. The more consistent an individual's response waveform across tasks, the more reliably his or her performance can be monitored

under changing circumstances.

To date, we have run a sample of 20 young adults in four tasks which were chosen to produce P300s which then could be evaluated for consistency across the four tasks.

Task 1 required subjects to count the number of occurrences of one of two equiprobable tones which differed in pitch. In general, P300 is larger for the counted than for the uncounted tones in this paradigm.

In task 2, the subject pressed a button with the left thumb when one tone pitch occurred and a different button with the other thumb when the other pitch occurred. The tones differed in probability (20% and 80%). The rare tone typically elicits a larger P300 than the frequent tone.

Task 3 was somewhat different. Only one tone pitch was used. On 10% of the trials, the tone was not presented. The subject was to count the number of "omitted stimulus" trials. P300 is usually larger on such trials.

Task 4 was a visual analog of task 2. Male names were presented on 20% of the trials, female names on 80%. Again, the rare class of stimuli should elicit a larger P300.

The P300 component in each average ERP was then scored with a vector filter technique (Gratton et al, in preparation, see 4.1) which detects the P300 by evaluating the distribution of the ERP.

Subjects were able to perform the tasks well, with few errors. A considerable amount of cross-task consistency in P300 amplitude and in overall wave shape from visual inspection of the data. This impression was statistically confirmed by a significant Kendall coefficient of concordance ($W=.501$, $p<.006$) for P300 amplitude.

Thus, our data indicate that young adults do show consistency across tasks. We plan to expand our subject sample and also to employ additional methods of quantifying consistency in order to verify this conclusion.

APPENDIX A

Items referred to in the Introduction.

1. Fabiani, M., & Donchin, E. P300 and memory: Individual differences in the von Restorff effect. Cognitive Psychology, in press.

2a. Fabiani, M., Karis, D., Coles, M. G. H., & Donchin, E. P300 and recall in an incidental memory paradigm. Psychophysiology, 1983, 20, 439. (Abstract)

2b. Handout

3. Coles, M. G. H., Gratton, G., Dupree, D., Bashore, T. R., Eriksen, C. W., & Donchin, E. P300 and response accuracy: An analysis using response bias and error titration. Biological Psychology, in press. (Abstract)

4. Gratton, G., Coles, M. G. H., Bashore, T. R., Eriksen, C. W., & Donchin, E. An ERP/EMG/RT approach to the continuous flow model of cognitive processes. Paper presented at the Evoked Potential International Congress (EPIC) meeting, Florence, 1983.

5a. Kramer, A., Fisk, A., & Schneider, W. P300, practice and consistency in visual search. Psychophysiology, 1983, 20, 453-454. (Abstract)

5b. Handout

6a. Gratton, G., Coles, M. G. H., Donchin, E. Filtering for spatial distribution: A new approach (Vector Filter). Psychophysiology, 1983, 20, 443-444. (Abstract)

6b. Handout

7. Mane, A. M., Coles, M. G. H., Wickens, C. D., & Donchin, E. The use of the additive factors methodology in the analysis of skill. Proceedings of the Human Factors Society 27th Annual Meeting, 1983, 407-411.

8a. Mane, A., Sirevaag, E., Coles, M. G. H., & Donchin, E. ERPs and performance under stress conditions. Psychophysiology, 1983, 20, 458. (Abstract)

8b. Handout

APPENDIX A1

Cognitive Psychology
in press

"P300" and Memory:
Individual Differences in the von Restorff Effect

Demetrios Karis, Monica Fabiani
and Emanuel Donchin

Cognitive Psychophysiology Laboratory
University of Illinois
Champaign, Illinois

Running head: P300 and Memory

Send proofs to: Dr. Emanuel Donchin
Psychology Department
University of Illinois
603 East Daniel
Champaign, Illinois 61820

Abstract

Event related brain potentials (ERPs) were elicited by words in a free recall paradigm that included a novel item. The P300 component of the ERP is elicited by novel, task-relevant events, and we tested the hypothesis that P300 is a manifestation of the cognitive processing invoked during "context updating." If the degree to which current representations in working memory need revision is related to P300 amplitude, then the P300 elicited by a given item should be related to the ability to recall that item on a subsequent test. Forty lists were presented to 12 subjects in each of two sessions. The lists were 15 words long, and one word, in position 6 through 10, was "isolated" by changing its size. Most subjects recalled these isolated words more often than other words in the same positions (von Restorff effect), and these words also elicited larger P300s than other words. Analysis of variance on the component scores from a principal components analysis revealed that words recalled had a larger amplitude P300 (on initial presentation) than words not recalled. Striking individual differences emerged, and there were strong relationships between the von Restorff effect, overall recall performance, mnemonic strategies, and the association between components of the ERP and recall performance. The overall recall performance of subjects who reported simple (rote) mnemonic strategies was low, but they showed a high von Restorff effect. For these subjects the amplitude of the P300 elicited by words during initial presentation predicted later recall. In contrast, subjects who reported complex mnemonic strategies remembered a high percentage of words and did not show a von Restorff effect. For these subjects P300 did not

predict later recall, although a later "slow wave" component of the ERP did. The initial response to isolated items was the same for all subjects (a large P300), and all subjects recognized the isolates faster than other words in a recognition test given at the end of each session. The subjects in whom P300 did not predict recall reported mnemonic strategies that involved organizing the material. These strategies continue long after the time period reflected by P300 (600 msec). Because they were so effective they may have overshadowed the relationship between P300 and recall, which is based on the initial encoding of an event. Our interpretations were further confirmed and clarified from data obtained in a final grand recall and in the recognition test.

DESCRIPTORS: Event-related potentials, P300, memory, individual differences, cognitive processing, strategies, von Restorff

"P300" and Memory: Individual Differences in the von Restorff Effect

Demetrios Karis, Monica Fabiani, & Emanuel Donchin

The label von Restorff, or isolation effect, refers to the enhanced learning of an "isolated" item (von Restorff, 1933). It was discovered within a context of the Gestalt psychologists' attempt to develop a field theory of recall, based on principles of interaction in perception (Koffka, 1935; Kohler, 1940). The effect is very robust and has been replicated repeatedly (Cimbalo, 1978; Wallace, 1965). When one item in a list is distinctly different from the others (e.g., because of color, size, meaning, or class) the probability that it will be recalled increases. Since isolated items are both novel and task-relevant there is a striking similarity between their attributes and the attributes of stimuli that elicit the P300 component of the human event-related brain potential (ERP). In ERP experiments novel, task-relevant, events elicit a positive potential with a latency to the peak of at least 300 milliseconds following the eliciting stimulus. This component of the ERP, commonly called P300, is a manifestation at the scalp of intracranial activity involved in cognitive processing (Donchin, 1979, 1981).[1] The data currently available on the conditions under which P300 is elicited suggest that P300 reflects processes invoked when there is a need for "context updating"; that is, when there is a need to revise the current representations in working memory (Donchin, 1981; Nageishi & Shimokochi, 1980).[2]

It is well established that P300 is elicited by unexpected events, and that the lower the subjective probability of an event the larger will be the P300 it elicits (Duncan-Johnson & Donchin, 1977). However, this strong effect of probability is restricted to task-relevant events and is tempered by the time interval between successive occurrences of the eliciting events, suggesting that P300 is sensitive to the strength of a decaying memory representation. These factors were combined by Squires, Wickens, Squires, and Donchin (1976) to form a predictive model that described the effects of probability on P300. They demonstrated that the P300 amplitude elicited by an event is affected by the sequence of preceeding events.[3] The model that accounted successfully for the data assumed that the strength of the memory trace decayed as an exponential function of the time that had passed since the last presentation of the stimulus. Heffley (1981) directly investigated the effects of varying the interstimulus interval (ISI). In his experiments he used ISIs of 6, 3, and 1.3 seconds and found that target probability had no effect on P300 amplitude at an ISI of six seconds. At this ISI all stimuli elicited an equally large P300. It was only at the shortest ISI of 1.3 seconds that the low probability stimuli elicited P300s larger than those elicited by the high probability stimuli. P300 amplitude to a task-relevant event is thus strongly influenced by the intervals between repetitions of that event. Presumably, the time period between repetitions of task-relevant events influences the strength of the representation in working memory. If the representation is weak, more updating must follow target presentation. At long ISIs all relevant events, regardless of probability, elicit large P300s. At short ISIs, on the other hand, only

rare targets need updating upon presentation, because frequent events are likely to occur while their previous representation is still held in working memory. At short ISIs then, only rare targets will elicit large P300s. Note that the attenuation of the probability effect at long ISIs is due to an increase in the amplitude of the P300 elicited by the frequent events, rather than to a decrease of the amplitude of the response to the rare events. This observation is consistent with the view that P300 amplitude is related to the amount of updating which is required by a task-relevant event. Fitzgerald and Picton (1981) provide further support for this interpretation. Using a simple auditory paradigm (count rare tones) Fitzgerald and Picton found that P300 amplitude elicited by the counted tones increased as the ISI was increased. In their experiment sequential probability was constant, p(target) = .20, but as the ISI was manipulated the temporal density of the targets (called by Fitzgerald and Picton the "temporal probability") varied. At the shortest ISIs (250 and 500 msec) the target tone was occuring so frequently that the previous target could still be in working memory (at 250 msec there was a target every 1.25 seconds, on the average, while at 500 msec one occurred every 2.5 seconds). P300 was small at these ISIs. The largest increase in P300 amplitude was between ISIs of 500 msec and 2 seconds, an interval during which temporal probability increased to 1 target every 10 seconds. Since targets would be unlikely to remain in working memory for these intervals they would require updating, and P300 would increase.

In considering the function of the process manifested by P300 it is useful to note that the elicitation of this process on a particular trial is

not necessarily critical for the execution of responses on that trial. For example, Kutas, McCarthy, and Donchin (1977) have shown that the relationship between overt responses and the P300 depends on the subject's strategy. The P300 appears to be elicited with a latency that depends on the time necessary for stimulus evaluation, whether or not the subject has already responded to the stimulus. Donchin, Ritter, and McCallum (1978) interpreted these, and similar data, to imply that the P300 process is invoked in the service of future-oriented activities related to the subject's subsequent strategies, rather than to the immediate "tactical" responses to the stimuli. One possibility is that the P300 is a manifestation of processes that maintain an accurate environmental model, or schema, by continually revising this model according to the most recent, useful data acquired by the nervous system.[4] The schema in this context is viewed as a large and complex map representing all the available data about the environment (Donchin, 1981). When there is a need, the schema is revised by the incorporation of incoming data. This updating process is manifested by the P300. Theories of human (Sokolov, 1963, 1969, 1975) and animal (Wagner, 1976) memory have also argued that a short term memory is used to maintain an internal model of a dynamic environment, and that deviations from this internal model require an updating process.

It is obvious that any adequate model or schema must represent more than just the most likely outcome in any situation. Regular, but rare events cannot be totally unexpected. The system, however, will not be "primed" for these rare events. It is possible that the schema is "activated" to different degrees. The aspects that are central to the

current task are most active and they constitute the "working" memory. When less central aspects of the schema must be activated to allow processing of novel, or rare, events other segments of the schema are scrolled into working memory. The process whereby the working memory is modified in response to environmental events is manifested by the P300. We assume that the amplitude of the P300 is proportional to the amount of change that was required in working memory by the environmental events. Further, we assume that the schema is continually being modified, some times gradually, at other times suddenly, and P300 reflects the nature of this process. The initial creation of structure is hypothesized to occur in working memory, and updating processes are also likely to occur there. As Broadbent (1981) writes, "The frequency with which an event has occurred can of course be counted in the nervous system without entering into a working memory. It is only the formation of fragments, the creation of structure, which needs the holding of temporary representations. This in turn means that structure is created only from selected aspects of the environment, those that have been encoded in working memory" (p. 22).

P300 is certainly not necessary for memory, but memory for events that elicit a P300 will, in general, be better than for events that do not. In previous research using a recognition paradigm we found that words in the study phase elicited only small P300s, if any at all (Karis, Bashore, Fabiani, & Donchin, 1982). Since some of these words were both recognized and later recalled, we must assume either that the updating process was in some way diminished, and therefore invisible to our scalp electrodes, or else that some other processes were involved. We do not yet have enough

information to choose between these two possibilities. If the amplitude of the P300 is proportional to the degree of memory updating, then it is likely that the P300 amplitude elicited by a given item will be proportional to the likelihood that this item will be recalled in a subsequent memory test. Therefore, "isolated" items, in the von Restorff sense, that are recalled should elicit a larger P300 than isolated items that are not recalled. We report here a study designed to test the hypothesis that the larger the P300 elicited by an isolated item the more likely it is to be recalled. To the extent that other words elicit a P300, they too should show this relationship.

METHOD

Subjects

Twelve right handed female subjects were run in two sessions that were separated by at least one week (range = 7 to 15 days, mode = 7 days). All were undergraduate students at the University of Illinois (age range 18 to 21). They were paid $3.00 per hour, with a $5.00 bonus when they completed the second session.

Word Lists

Two word lists were constructed for each subject according to the following rules: for each session and subject a computer program selected words at random from one of two longer lists (one per session) composed of all the actual words with 3 to 6 letters in Toglia and Battig (1978). Each word was presented no more than once to each subject.

Words could appear in one of three sizes: small, medium or large.

Small words were formed with letters of 7mm x 7mm and ranged in length from 21mm to 42mm (visual angle = 1.35 to 2.70 degrees). Medium words were formed by 12mm X 12mm letters (word length 36mm to 72mm, visual angle 2.25 to 4.50 degrees). Large words were formed by 20mm X 20mm letters (word length, 60mm to 120mm, visual angle 3.75 to 7.50 degrees). Size differences among these letter sizes were easily discriminable.

Data Collection

Burden Ag-AgCl electrodes were affixed with collodion along the midline of the scalp at frontal, central, and parietal sites (Fz, Cz, and Pz according to the 10/20 International System; Jasper, 1958) and with adhesive collars to each mastoid. Ag-AgCl Beckman Biopotential electrodes were used as ground and electrooculogram (EOG) electrodes. The subject was grounded on the forehead, and sub- and supra-orbital electrodes were used to record the EOG. Linked mastoids were used as the reference sites. Electrode impedance did not exceed 10 kOhm. The EEG was amplified with Van Gogh Model 50000 amplifiers (time constant 10 seconds, upper half-amplitude frequency 35 Hz, 3dB/octave roll-off) and was digitized at the rate of 100 samples/sec for 1280 msec, beginning 100 msec prior to stimulus onset.

All aspects of experimental control and data collection were controlled by a PDP-11/40 computer system interfaced with an Imlac graphics processor (Donchin & Heffley, 1975). Average waveforms and the single-trial records were monitored on-line using a DEC VT-11 display processor. Eye movement artifacts were corrected off-line using a procedure described in Gratton, Coles, and Donchin (1983a).

PROCEDURE

The subject was seated in an air conditioned unshielded room in front of a Hewlett Packard (HP) CRT display (#1310A). The recording and control apparatus were located in an adjacent room. Each session comprised four tasks: free recall, a counting task ("oddball" paradigm), a final grand recall, and a recognition test (see Figure 1). In the first session the

----

Insert Figure 1 About Here

----

final grand recall and the recognition test were unexpected. In the second session subjects were told that there would be a grand recall and recognition, as in the first session, but that they should not worry about these, as the free recall was of primary importance. The EEG data were acquired whenever a stimulus word was presented on the HP screen (i.e., during the free recall, oddball, and recognition phases). The stimulus duration in all tasks was 200 msecs with a 2 second ISI.

A. Free recall

Forty lists of 15 words each were presented to the subject during each session. Words in each list were presented sequentially with a 2 sec interval between words. In thirty out of the 40 lists one of the words, in the sixth through the tenth position, was an "isolate". The isolation was achieved by displaying the word in either larger, or smaller, characters than those used to display the other words (see "Word Lists", above). The other ten "control" lists did not include an isolated word. The specific location of the isolate in each list was randomly selected, as was the order

of presentation of experimental and control lists.

The subject was instructed to memorize as many words as she could, and was given a clipboard with 40 sheets (one per list) on which to write the words after each list was completed. A 7 second pause was interposed at the end of each list, during which she was instructed to turn the page. At the end of the pause a small light attached to the clipboard was turned on, signaling the subject to pick up the pen and start writing. Removal of the pen from its holder activated a switch monitored by the experimenter, so that the subject could not begin writing prematurely. Fifty seconds were provided for the free recall, and all subjects reported that this interval was sufficient. The writing light was then turned off to indicate that the recall period had ended. After a verbal warning ("ready?") from the experimenter (via an intercom), another list was presented. The subject was allowed to rest after every ten lists. Two practice lists were given to the subjects at the beginning of their first session. One was an experimental list containing a small isolate, the other was a control list. After the experimental list subjects were asked if they had noticed that one word was smaller than the others, and were told that occasionally a word would appear larger or smaller, but that they should attend to all the words, and ignore size differences between words.

B. Oddball

An "oddball" task was presented after the free recall.[5] It served to fill the interval between free recall and grand recall. It also served to provide a record of the ERPs in a paradigm comparable to that used in other studies. The subjects were presented with a series composed of the word

"count" presented 100 times. On 20 trials the characters were either larger or smaller in size than the other 80 trials. The size of the rare stimulus was counterbalanced across subjects. Subjects were instructed to count the rare stimuli (subvocally) and to report the running total at the end. This total was usually correct, and was always within one of the actual number.

C. Grand recall

After the oddball, the subject was asked to write down all the words she could remember from any of the lists presented during the free recall phase. Ten minutes were provided for this task.

D. Recognition

Finally, the subject was presented with a sequence of 120 words all displayed at the same size. Sixty of these words (50%) had already been presented in the free recall phase. Of these, thirty had been the isolated words (all the isolates were included), while the other thirty included one word from each of the experimental lists, with the limitation that the four words surrounding the isolated word (two before and two after) not be chosen. The other 60 words (50%) were new words chosen from the same master list used to generate the free recall lists. The subject was instructed to press one of two buttons (using her thumbs) to indicate whether the word presented was a word she had seen before or a new word. (No discrimination was required between isolates and the other old words.) Subjects were instructed to be as quick as possible without sacrificing accuracy. Response hands were counterbalanced across subjects.

E. Debriefing

At the end of each session the subject was asked about the strategies she used in memorizing the words during the free recall phase. These descriptions were subsequently rated, as to the strategy used, by nine undergraduates who had not participated in the experiment, and who were not aware of the purpose of the study. The rating task will be described in detail below.

RESULTS

Analysis of Recall and Recognition

A. Free and Grand Recall

We computed two indices to summarize the subjects' performance in the free recall task: a measure of the von Restorff effect (Von Restorff Index, or VRI) and an index of overall recall performance (P). Both indices were computed using the words recalled by the subject in the free recall. Only words originally presented in position 6 through 10 were used to calculate the VRI (in order to match the positions of isolates and non-isolates). In the computation of the overall recall performance all the words were used. VRI and P were computed as follows:

VRI = percentage of isolated words recalled (position 6-10) - percentage of non-isolated words recalled (position 6-10)

P = overall percentage of words recalled from all positions (isolates and non-isolates)

Since there were no systematic differences in the recall of non-isolated words coming from control and experimental lists ($F = 3.17$; df = 2,18; $p > .05$), non-isolated words from both experimental and control lists were used to compute the VRI. Analogous indices (VRI and P) were also computed for the grand recall phase, using the words recalled by the subject in the grand recall.

The values of these two indices (VRI and P), calculated from all 80 lists, are plotted for all subjects in Figure 2. It is clear that subjects

---

Insert Figure 2 About Here

---

differed in their performance. There appear to be three clusters of subjects according to the VRI assessed in the free recall period. Group 1 is composed of subjects who showed a high von Restorff effect (the upper quartile of VRI distribution), in group 2 are subjects with a medium VRI (the intermediate quartiles), and group 3 is composed of subjects with a low VRI (the lowest quartile). VRI and P for each subject (from both free recall and grand recall), as well as overall and group means and SDs are presented in Table 1.[6] The term "improvement" in Table 1 refers to the

---

insert Table 1 About Here

---

change in percentage of recalled items from session 1 to session 2. It is noteworthy that the group subdivisions correspond to actual gaps in the VRI

Since there were no systematic differences in the recall of non-isolated words coming from control and experimental lists ($F = 3.17$; df = 2,18; $p > .05$), non-isolated words from both experimental and control lists were used to compute the VRI. Analogous indices (VRI and P) were also computed for the grand recall phase, using the words recalled by the subject in the grand recall.

The values of these two indices (VRI and P), calculated from all 80 lists, are plotted for all subjects in Figure 2. It is clear that subjects

---

Insert Figure 2 About Here

---

differed in their performance. There appear to be three clusters of subjects according to the VRI assessed in the free recall period. Group 1 is composed of subjects who showed a high von Restorff effect (the upper quartile of VRI distribution), in group 2 are subjects with a medium VRI (the intermediate quartiles), and group 3 is composed of subjects with a low VRI (the lowest quartile). VRI and P for each subject (from both free recall and grand recall), as well as overall and group means and SDs are presented in Table 1.[6] The term "improvement" in Table 1 refers to the

---

insert Table 1 About Here

---

change in percentage of recalled items from session 1 to session 2. It is noteworthy that the group subdivisions correspond to actual gaps in the VRI

distribution. It is also evident that VRI and P are inversely related: the lower the subject's von Restorff effect the higher her performance. This relation holds for the grand recall as well, even though both VRI and P decrease. Table 2 presents the Pearson correlation coefficients between VRI and P in both free recall and grand recall.[7] An analysis of variance (ANOVA)

----

insert Table 2 About Here

----

on the same data reveals that the three groups differ significantly from one another ($p < .05$) with respect to both the von Restorff effect and their overall performance in both free and grand recall (free recall: for VRI, $F = 71.93$, df = 2,9; for P, $F = 18.60$, df = 2,9; grand recall: for VRI, $F = 7.3$, df = 2,9; for P, $F = 4.74$, df = 2,9). Group 1 includes subjects who show a high von Restorff effect but are overall poor memorizers, group 3 consists of people who show no von Restorff effect but who are very good memorizers, and group 2 subjects shows an intermediate level of performance on both dimensions.

Serial position curves (from the free recall) are shown in Figure 3 for each group. They show the percentages recalled for words from each of the

----

insert Figure 3 About Here

----

15 list positions. Data for isolated and non-isolated words are plotted separately. Note first that all three groups show a "primacy", as well as a "recency" effect. The magnitude of the von Restorff effect in each group

can be seen in Figure 3, as it is represented by the elevation in recall of the isolated items (the triangles) relative to the rest of the curve. Clearly, the von Restorff effect is largest in group 1, moderate in group 2, and absent in group 3.[8]

B. Recognition

For the recognition test, median reaction times (RT) for words correctly recognized and error rates (ER) were computed for each subject and each class of words (isolates, non-isolates and new). The same subject grouping described above was maintained for the analysis of recognition performance. Group means and standard deviations (SDs) are shown in Table 3. An overall effect of word type on RT's was found: in all groups subjects

---

insert Table 3 About Here

---

respond faster to isolates than to non-isolates, and the slowest RT's correspond to the new words ($F = 65.00$; df = 2,18; $p < .05$). Subjects in group 3 have longer RT's to the new words than subjects in the other two groups (group x word interaction: $F = 7.35$; df = 4,18; $p < .05$). It is important to remember that subjects were asked only to indicate whether a word was "old" or "new" and that the correct response for both isolates and non-isolates was the same (old). No significant correlation between RT and ER was found ($r = .15$; $p > .05$), and there were no differences among groups in error rates.

C. Session Effects

The data above are combined across both sessions. When the pattern of results was examined separately for each session, we found no significant interactions between sessions and groups on other variables, with the exception of performance. Subjects in all groups improved from the first to the second session in both free recall and grand recall, with group 3 improving the most. The subject's improvement (I) was calculated by subtracting the performance in session one from the performance in session two.

Degree of improvement in both free and grand recall is reported in Table 1 for each subject. Improvement in free recall is plotted against the von Restorff index in Figure 4. Subjects in the three groups show different degrees of improvement: subjects in group 1 improve, in the free recall task, less than subjects in group 2 and 3 (main effect of group: $\underline{F} = 7.86$; df = 2,9). Improvement is also correlated with VRI and P, and correlations

---

insert Figure 4 About Here

---

are reported in Table 2.

To calculate split-half reliability for performance and the von Restorff index we correlated the relevant scores obtained in the two sessions and applied the Spearman-Brown formula. For free recall the reliabilities were .96 for performance and .62 for the von Restorff index. The von Restorff index is based on far fewer trials than performance; thus, it is not surprising that reliability is smaller. The reliabilities of the

VRI and performance calculated from the grand recall were .74 and .79, respectively. Given that the two sessions were at least one week apart, these correlations suggest that these individual differences were stable.

D. Strategy Reports

Subjects' reports about the strategies they used to memorize the words were rated, blindly, by 9 undergraduate students who were paid to serve as judges. They were instructed to rank order the strategies from the most simple (rote) to the most complex (elaborative). The rote strategies were defined in the instructions as "simple strategies, mainly involving repeating each word, or group of words, over and over". The elaborative strategies were defined as "complex strategies, mainly involving combining the words into stories, or producing complex images or sentences". An example of a rote strategy is given by this subject from group 1: "...I repeated the words in a row. I also tried to repeat each word three times..." (mean rank = 1.7). A subject from group 3 gave this report: "...I tried to connect words into a story or a picture. I tried to make the story or the picture ridiculous..." (mean rank = 11.8). Inter-judge reliability (as measured by Cronbach's Alpha) was .98, and the correlation between the VRI and mean rank given to the subject's strategy was -.57, (p < .05). A high von Restorff effect is associated with a low rank (which indicates rote strategies). The mean ranks for groups one,two, and three were 3.6, 6.5, and 9.2. It turns out, then, that the three groups differed markedly in their choice of encoding strategies. The difference between groups 1 and 3 are particularly striking. The subjects in group 1 are primarily rote memorizers, while group 3 subjects are "elaborators" or

organizers. Subjects in group 2 take an intermediate position.

DISCUSSION

Striking individual differences emerged on all measures, and subjects were placed into three distinctly different groups based on their von Restorff index from the free recall. In group 1 subjects' overall performance was low, but "isolating" a word by changing its size increased recall dramatically (high von Restorff effect). These subjects reported using primarily rote strategies and did not improve across sessions. At the other extreme, subjects in group 3 exhibited high overall performance, and there was no effect of isolation on recall. These subjects reported complex, associative strategies and improved significantly across sessions. Subjects in group 2 were intermediate on all measures (overall performance, the von Restorff effect, and improvement) and reported using a variety of mnemonic strategies. In the grand recall, where recall from all 40 lists was requested, performance was generally reduced for all subjects, but the group differences remained. The von Restorff effect and performance calculated from the grand recall were significantly correlated with these same measures in the free recall; e.g., subjects who showed a strong von Restorff effect in the free recall also tended to produce a large effect in the grand recall. This is the first time, in our knowledge, that individual differences in the von Restorff effect have been studied. We will discuss these differences in detail below, in conjunction with the ERP results.

A Note on the Use of Strategy Reports

In the last few years there has been much discussion on the merits of

using verbal reports as data in behavioral experiments (Ericsson & Simon, 1980; Kellogg, 1982; Nisbett & Wilson, 1977; Smith & Miller, 1978; White, 1980). We agree with Morris (1981) that a distinction should be made between "strategy reports", which describe "consciously chosen strategies", and "self-hypotheses", in which subjects try to "describe the causes of their behavior" (p.465). All verbal reports, of course, must be assessed carefully in the context of the experimental demands. However, when experimenters ask for strategy reports, and not for self-hypotheses, then introspective reports may lead to insights about cognitive activity, and help in understanding individual differences.[9]

We found that the strategy reports were useful in elucidating the differences between the three groups, and in understanding the relationship between overall performance, the magnitude of the von Restorff effect, improvement across sessions, and the ERP results. We had no a priori hypotheses about the relationship between these variables; on the contrary, we expected all subjects to exhibit a strong von Restorff effect. It is also important that the experimenter (M.F.) was not aware, at the time she debriefed the subjects, of the von Restorff index or the general recall ability of the subjects. These indices were computed after the debriefing.

ERP RESULTS

A. Free Recall

As above, data for each subject were combined across sessions. EEG records related to each word were sorted for averaging by word type (isolates, non-isolates in experimental lists, and control words), by position (position 6-10, other positions), and by subsequent recall

(recalled, not-recalled). Our main interest was in the comparisons between isolates and other words in the same position (6 through 10), and the analyses below will be restricted to words in these positions.

Average waveforms at Pz for isolates, non-isolates, and words from the control lists (control words) are shown in Figure 5 for each of the subjects. As can be expected, large P300s were elicited by the isolates, while only small P300s (or none at all) were elicited by non-isolates and control words. This relationship held for all subjects. This result is

---

insert Figure 5 About Here

---

consistent with the general observation that task relevant, distinct stimuli elicit a larger P300 than do companion stimuli that are common. Given that isolates did elicit P300s we proceeded to determine if there was any relationship between the amplitude of the P300 and subsequent recall. All isolates, for each subject, were therefore sorted into two categories, isolates that were, and those that were not, recalled in the free recall test immediately following list presentation. For each of the subjects we computed two ERP averages: one elicited by subsequently recalled isolates, and one by subsequently unrecalled isolates. This procedure was repeated for the other word types. We then averaged these ERPs across subjects according to the three groups of subjects identified above.

Average waveforms at Pz for the 3 groups and the 3 classes of words are presented in Figure 6. We note two primary aspects of these results. First, that it was only the isolates that elicited large P300s. Second,

that only in group 1 can a large difference in P300 between isolates recalled and not recalled be observed. Figure 7 shows group waveforms

---

insert Figure 6 About Here

---

elicited by the isolates at the 3 electrode sites (Fz, Cz, Pz). For group 1, the difference in P300 elicited by the recalled and non recalled items is prominent. It extends across electrodes in the typically parieto-maximal P300 distribution. However, note that differences also appear between ERPs associated with recalled and unrecalled isolates in group 3, and to a lesser extent in group 2. However, these differences are associated with a slow wave component with a frontal maximum that follows the P300.

---

insert Figure 7 About Here

It is possible that a single subject's data can dominate an ERP average when the group size is small. This did not happen here. In group 1 all three subjects showed the effect (P300 larger for words recalled), while in group 3 one subject showed the effect and two exhibited a slight reversal. In group 2 there was variability, but no subject showed an effect as strong as any of the subjects in group 1.

These impressions were corroborated by means of a a Principal Component Analysis (PCA; Donchin & Heffley, 1979) performed on all the average EEG records associated with words in position 6 through 10. Two hundred and sixteen waveforms were entered in the PCA (12 subjects x 3 words (isolates,

non-isolates, controls) x 2 memory levels (recalled, not recalled) x 3 electrodes). Four components (explaining 92% of the variance) were rotated using a Varimax rotation procedure. Component loadings are shown in Figure 8. The first three components were labelled, according to their latency and

---

insert Figure 8 About Here

---

their scalp distribution, as P300 (component 1), "frontal positive slow wave" (component 2) and N200 (component 3). Component scores for isolates are presented in Figure 9 for the first two components.

---

insert Figure 9 About Here

An analysis of variance was applied to the PCA component scores to test the differences in amplitude of each component over different experimental conditions. A repeated measures design with a nesting factor (group) and unequal Ns was used (ALICE statistical package, program "UNEN", Grubin, Bauer and Walker, 1976).

Significant results ($p < .05$) obtained for the first two components are described below. The complete ANOVA results are presented in Table 4. The

---

insert Table 4 About Here

---

only significant effect on the third component, N200, was a word by electrode interaction (the parietal negativity was larger for the isolates;

F = 8.64, df = 4,36). There were no significant effects associated with the fourth component.

Component 1: "P300"

We label component 1 "P300" because its peak latency (520 msec; see footnote 1) and its scalp distribution are characteristic of the P300 component of the ERP. Amplitude values are positive at the three electrode locations, with Pz more positive than Cz, and Cz more positive than Fz (main effect of electrode: F = 37.14; df = 2,18).

Isolated words show a larger P300 than control and experimental words (main effect of word: F = 20.73; df = 2,18), and words that are recalled (regardless of their type) show a larger P300 than words not recalled (main effect of memory: F = 14.09; df = 1,9).

Isolated words show larger positivity at Pz than control and experimental words (electrode x word interaction: F = 42.90; df = 4,36). Group 1 subjects (poor memorizers with a high VRI) show a larger P300 for words recalled than not recalled. The amplitude difference is smaller for group 2 and virtually absent for group 3 (group x memory interaction: F = 5.78; df = 2,9).

Finally, the largest amplitude difference for isolates recalled in comparison with the isolates not recalled is observed at Pz for group 1 (group x word x memory x electrode interaction: F = 3.93; df = 8,36).

Component 2: "Frontal-positive slow wave"

This component appears at 540 msec. and slowly increases until the end of the epoch. It is more positive frontally than centrally and parietally.[10]

Group 3 subjects (good memorizers with a low VRI) show more evidence of this component than the other two groups (main effect of group: $F = 5.71$; df = 2,9). This component is also more evident for the isolates than for the other types of words (main effect of word: $F = 3.83$; df = 2,18) and more evident for words recalled than not recalled (main effect of memory: $F = 10.78$; df = 1,9).

Further, this component is frontally more positive for isolates than for non-isolates and controls, which display a flat distribution (electrode x word interaction: $F = 7.39$; df = 4,36). It is also frontally more positive for words recalled than not recalled (electrode x memory interaction: $F = 6.64$; df = 2,18). Finally, this component shows the largest frontal positivity to isolates recalled by subjects of group 3 (electrode x word x memory x group interaction: $F = 2.37$; df = 8,36).

In summary, statistical analysis of the free recall waveforms shows that P300 is indeed largest when elicited by isolated items, and that for some subjects the larger the P300 elicited by a word the more likely is it to be recalled. The association between recall and P300 is most prominent for subjects in group 1. It is small, if not absent, in the other two groups. The frontal-positive slow wave component is also related to isolation of stimuli and to recall. However, for this component the effect is most evident in subjects of group 3.

It is important to determine if groups 1 and 3 differed in the manner in which they reacted to the isolates. If the two groups differed in the distribution of P300 amplitudes across trials such differences may account for the relationship between P300 and recall. We note that there was no

difference among groups with respect to mean P300 amplitudes, as there was no significant group x word or group x word x electrode interaction on component 1. To examine, within subjects, the distribution of P300 amplitude elicited by isolates, we assessed the amplitude of P300 on individual trials by means of a set of weights obtained using a stepwise discriminant analysis (SWDA) (see Donchin, 1969b, and Horst & Donchin, 1980). This procedure was chosen in order to minimize the problems due to the low signal to noise ratio of single trials. A SWDA was performed for each subject, using rare and frequent trials from the oddball condition as the training set. The SWDA determines the subset of variables (chosen from all the variables entered in the analysis - in this case the timepoints) which best discriminates between the categories of interest (rare and frequent trials, in this case). The best discriminant linear combination of these variables results in a discriminant function that can be applied to classify any new group of individual trials. Given that the difference between rare and frequent trials in an oddball paradigm can be attributed (at least at first appoximation) to the difference in P300 amplitude, we used the discriminant function to assign a P300 amplitude score to each individual trial of the free recall. The distribution of amplitudes for groups 1 and 3 were compared by classifying scores into four amplitude categories and comparing the two resulting distributions using a chi square test. No statistically significant differences were observed between groups 1 and 3 ($\chi^2 = 5.59$, df = 3, $p > .05$). This result is consistent with the assumption that the initial processing of the isolates was similar for all subjects, and therefore that differences in the relationship between the

recall of an item and the amplitude of the P300 it elicits can be attributed to the subject's recall strategies (see below for details). The only significant effect on N200 is in accord with this argument. N200 is often related to a perceptual "mismatch detector" (see Naatanen & Gaillard, 1983). The isolates elicited larger N200s than the other words, and this effect held for all subjects, suggesting an equal processing related to the deviance of the isolates.

B. Grand Recall

No ERPs were recorded during the grand recall, but ERPs recorded during the free recall were sorted according to performance in the grand recall. Again we emphasize that the ERPs we examine are those elicited at the initial presentation of the word, though the trials are sorted according to subsequent recall performance. For each subject, we combined data from both sessions and averaged EEG records according to word type (isolate, non-isolate, control), word position (position 6-10; other positions), and subsequent recall (words not recalled during either the free recall or the grand recall; words recalled in the free recall but not in the grand recall; words recalled during both free recall and grand recall). Grand average waveforms over 12 subjects for the 3 classes of words at 3 electrode locations are presented in Figure 10. Only ERPs elicited by words in position 6-10 are included.

---

insert Figure 10 About Here

A PCA was performed on all the waveforms used to compute the grand average (12 subjects x 3 words x 3 memory levels x 3 electrodes = 324 waveforms). Four components were rotated using a Varimax procedure. Their latency and scalp distribution were quite similar to the components extracted in the PCA of the free recall (as expected, given the overlapping of the input waveforms).

An ANOVA was performed on the PCA component scores (as described above). Among the significant results ($p < .05$) was a main effect of memory for P300 ($F = 7.56$; df = 2,18). The largest amplitude P300 belongs to words that were recalled in both free recall and grand recall, while the smallest belongs to words never recalled.

C. Recognition

By combining the recognition and recall phases of our experiment we further tested the hypothesis that there is a graded relationship between "memory level" and P300 amplitude; i.e., the higher the probability of a word being subsequently recognized or recalled, the larger the P300 that will be elicited on its initial presentation. In general, recallable items can also be recognized (Watkins & Todres, 1978), although there are exceptions (Tulving, 1968; Tulving & Thomson, 1973; Watkins & Tulving, 1975). Given that recognition is usually much easier than recall, it is reasonable to expect a smaller P300 for words recognized but not recalled

than for words that were both recognized and recalled. In addition, we expect a larger P300 for words recalled in both free recalls than for words recalled in just the initial free recall. To test these hypotheses we reaveraged waveforms recorded when words were presented during the free recall phase on the basis of free recall, grand recall, and recognition performance. Words were sorted into four groups and four averages were computed: isolates neither recognized nor recalled, isolates recognized but not recalled, isolates recognized and recalled during the free recall, and isolates recognized and recalled in both the free recall and the grand recall. Other combinations, such as words recalled but not recognized, contained too few trials for analysis.

Grand average waveforms of all 12 subjects are shown in Figure 11 for Pz, and the expected gradations of P300 amplitude are visible in the

----

insert Figure 11 About Here

----

waveforms. No further statistical analysis was performed because too few trials per subject were averaged in each class. Note, however, that the a priori probability of obtaining the expected rank order of waveforms is 1/24 (Conditional probability = 1/4 x 1/3 x 1/2; $p < .05$).

EEG records recorded during the recognition test were also averaged for each subject, according to word type (isolates, non-isolates, new) and to recognition (correctly and incorrectly recognized). Grand averages over 12 subjects for 3 classes of words (correct and incorrect) at 3 electrode locations are shown in Figure 12.

---

insert Figure 12 About Here

A PCA was applied to the data (12 subjects x 3 words x 2 recognition levels x 3 electrodes = 216 waveforms) and four components were rotated using a varimax procedure. The components had latencies and scalp distributions very similar to the components extracted in the other PCAs.

There were significant results ($p < .05$) only for component 1 (P300). The P300 component (larger parietally - main effect of electrode: $F = 13.22$; df = 2,18) was larger for words correctly recognized than incorrectly recognized (main effect of recognition: $F = 6.78$; df = 1,9). Further, isolated and non-isolated words that were correctly recognized show larger P300s than new words correctly recognized (word x recognition interaction: $F = 11.89$; df = 2,18). These differences are easily visible in Figure 11, especially at Pz, where P300 is maximal.

Statistical "Problems" with PCA/ANOVA

The use of principal component analysis (PCA) followed by an analysis of variance (ANOVA) on the component scores is well established in ERP research (Coles, Gratton, Kramer, & Miller, in press; Curry et al., 1983; Donchin, 1969a; Donchin & Heffley, 1978). Several investigators have expressed some concern, however, about the appropriateness of using PCA on multiple records taken from the same individual (E. Hunt, 1980, personal communication, June 8, 1983; Wastell, 1981). Since the set of loadings used to extract the principal component scores are chosen in order to simplify

the variance/covariance matrix, the resulting ANOVA may be biased, increasing the probability of a type I error. We have taken two steps to support our assertions regarding the probability of type I error. First, we performed a "bootstrap" analysis in order to generate empirical distributions for the F values associated with our two most important interactions, and second, we performed ANOVAs on measures of P300 and slow wave amplitude that did not depend on the PCA. We utilized instead a new filtering technique known as "vector analysis".

1. The Bootstrap

Bootstrapping is a nonparametric method for estimating statistial accuracy from the data in a single sample (Diaconis & Efron, 1983; Efron, 1979; Efron & Gong, 1983). In general, the procedure generates an estimate of the distribution of the test statistic that does not depend on assumptions regarding the data. This is achieved by constructing "bootstrap samples" and performing many "bootstrap replications." Each sample is obtained by random sampling of cases from the pool of all available cases (with replacement), and the statistic of interest is then calculated for each such sample (the bootstrap replication). If this is done often enough, a distribution of the statistic is obtained. Efron and Gong (1983) describe this as "the substitution of raw computing power for theoretical analysis" (p. 36), and Efron and his associates present theoretical and empirical justification for this procedure.

We created 1004 bootstrap samples. Each was generated by picking a sample of N = 12 by random sampling (with replacement) from our pool of 12 subjects. (Of course, most of these bootstrap samples did not contain all 12

subjects, for some subjects were picked more than once.) The 12 subjects in each bootstrap sample were randomly divided into three groups of sizes 3,6, and 3 to correspond in size to our three groups. The P300 component scores (component 1) for each chosen subject were then entered into the same ANOVA on unequal Ns that is reported above. We examined only two values from each ANOVA, the group by memory interaction (GR x ME), and the group by word by memory by electrode interaction (GR x WO x ME x EL). These interactions indicate that the difference between P300 to recalled versus non recalled words varies across groups. The difference is largest in group 1, and smallest in group 3. In the 1004 bootstrap samples, only 30 times (30/1004 = .0299) was there an F value greater than the one we obtained for the GR x ME interaction (5.776). For the GR x WO x ME x EL interaction an F value greater than ours (3.926) appeared only 12 times (12/1004 = .0120)(see Efron & Gong, 1983, p. 42). We conclude, therefore, that the probability of a type I error is indeed lower than .05. Furthermore, when we examined the 42 bootstrap samples that generated larger F values than those we obtained, we found that the extreme groups in these samples (1 and 3) often contained a subject from our extreme groups that had been picked twice (i.e., of three subjects in one of the extreme groups, two were actually the same subject). This happened in 15 of the 30 samples in the GR x ME interaction, and 9 of the 12 samples in the GR x WO x ME x EL interaction. This always occurred in the samples that produced the largest F values.

2. Vector Filter

To provide an additional check of our results, we applied a new procedure (Vector Filter, Gratton, Coles, & Donchin, 1983b; Coles, Gratton, Kramer, & Miller, in press) that defines "target" components in terms of the scalp distribution of the voltage. All segments of an epoch that meet the scalp distribution criteria are considered to represent the component in question. This procedure provides an estimate of a specific component at each time point by combining the values obtained at all the electrodes. The electrodes are differentially weighted, in order to maximize the scores for a particular component (identified with a specific scalp distribution), but the weights are defined a priori. The procedure is mathematically equivalent to a rotation in the space defined by the electrode locations. Vector Filter yields a series of estimates of the "target" component, one for each timepoint. A traditional peak picking procedure may then be applied to the time series thus obtained. Note that by means of Vector Filter independent estimates of P300 and Frontal Positive Slow Wave may be obtained for each timepoint, since the set of weights used for the two components are orthogonal.

An ANOVA was applied to the estimates of P300 and Frontal Positive Slow Wave amplitude obtained with this procedure (the ANOVA design was the same used for analyzing the component scores). The results of this ANOVA were very similar to the results of the ANOVA applied to the component scores. All significant effects ($p < .05$) obtained via PCA/ANOVA were also significant in Vector Filter/ANOVA. (No interactions with electrodes could be examined, of course, because the vector filter combines data across

electrodes.) ANOVA on P300 peak amplitude (at Pz) obtained conventionally also yielded equivalent results. The concordance of all these analyses strongly confirms the original PCA/ANOVA results.

DISCUSSION

Our predictions were confirmed, and some unexpected findings emerged. The isolated words, which were both novel and task-relevant, elicited much larger P300s than the control or experimental words. Larger P300s were elicited by words that were recalled than by words not subsequently recalled. This relationship held for all word types, not just isolates. Most interesting, however, were the differences between the groups. Subjects in group 1 who showed a strong von Restorff effect, while their general recall was poor, tended to display a strong relationship between the amplitude of P300 and subsequent recall of the eliciting stimuli. Subjects in group 3 - who recalled well, but showed a very small von Restorff effect - displayed no relation between P300 amplitude and recall. Isolates, whether subsequently recalled or not, elicited a large P300. In group 2 we obtained an intermediate effect. The frontal-positive slow wave component was also sensitive to both isolation and the probability of recall, but group 3 (good memorizers with a low von Restorff index) exhibited more evidence of this component than the other two groups. We will discuss these group differences in P300 and the frontal-slow wave component in the General Discussion.

The relation of memory processes to the entities manifested on the scalp by P300 was demonstrated when we reaveraged the ERPs collected during the free recall phase on the basis of the additional information collected

during the grand recall and recognition. Our assumption was that when we combined our three measures of memory (free recall, grand recall, and recognition) we would have a more sensitive index. A word recalled not only in the immediate free recall, but also 30 minutes later in the grand recall, should have a "stronger" representation in memory than a word recalled only in the first free recall. The P300 elicited by the initial presentation of a word was correlated with memory strength defined on the basis of the free recall tests. P300 was larger when a word was recalled in both free recalls than when it was recalled only during the first. Similarly, there was a graded change in P300 when recognition performance was added to the two free recalls, although this analysis could be performed only on isolates, and not enough trials were collected for statistical analysis. Nevertheless, P300 to the isolates, in order of increasing amplitude, was as expected: 1. neither recognized nor recalled, 2. recognized but not recalled, 3. recognized, and recalled only during the free recall, and not the grand recall, and 4. recognized and recalled in both free and grand recalls.

Recognition

We also recorded ERPs while words were presented during the recognition phase. In accord with our previous studies (Karis et al., 1982) we found that "old" items (both isolates and nonisolates) elicited larger P300s than "new" items (for correct responses), and that P300 was larger to words correctly recognized than to words incorrectly recognized. The larger P300 to old words is probably a combination of factors, including the "target effect" and confidence in the decision process. P300 amplitude increases

with the confidence with which a decision is made (Paul & Sutton, 1972; Hillyard, Squires, Bauer, & Lindsay, 1971; Squires, Squires, & Hillyard, 1975a; Squires, Squires, & Hillyard, 1975b), and targets elicit larger P300s than nontargets (Duncan-Johnson & Donchin, 1977).

Memory and ERPs

Most studies which have focused on the relationship between memory processes and ERPs have used a variety of recognition paradigms (Warren, 1980; Stanny & Elfner, 1980; Parasuram, 1980; Parasuraman, Richer, & Beatty, 1982) or the Sternberg task (Roth, Kopell, Tinklenberg, Darley, Sikora, & Vesecky, 1975; Marsh, 1975; Gomer, Spicuzza, & O'Donnell, 1976; Roth, Tinklenberg, & Kopell, 1977; Roth, Rothbart, & Kopell, 1978; Adam & Collins, 1978; Ford, Roth, Mohs, Hopkins, & Kopell, 1979). In the Sternberg paradigm ERPs are recorded to the probe stimuli, and the emphasis is usually on P300 latency (as an indication of stimulus evaluation time), rather than amplitude. Similarly, in most recognition experiments the emphasis is on the response to test stimuli, and whether or not there are ERP differences between old and new items (Warren, 1980), or between recognized and unrecognized stimuli in signal detection studies (Parasuraman, 1980; Parasuraman et al., 1982). In study-test paradigms ERPs are sometimes not recorded during the study phase (Stanny & Elfner, 1980), or are not averaged into separate classes on the basis of later performance (Warren, 1980). In previous work (Karis et al., 1982; see also Donchin, 1981) using a recognition paradigm we recorded ERPs during both the study and test phases, and in the free recall paradigm described here we also recorded ERPs during

the study phase. We then examined the relationship between ERPs elicited during the initial presentation of a word and subsequent performance during both free recall and recognition tests. An important aspect of our approach is that we sort trials according to the subject's performance and then average trials which are homogenous in some explicit respect. This procedure is, to our minds, of greater utility than the comparison of ERPs computed over blocks of trials that differ in some average performance score across the trials. Without an examination of performance on a trial by trial basis it is difficult, if not impossible, to make sense of the endogenous components of the ERP.

Both Chapman, McCrary, & Chapman (1978) and Sanquist, Rohrbaugh, Syndulko, & Lindsley (1980) report similar studies that used the design we advocate, in that they record ERPs during the initial presentation of an item, and relate these ERPs to subsequent memory tests. The research of Chapman's group, however, is difficult to interpret. Chapman et al.(1978; also reported in Chapman, McCrary, & Chapman, 1981) recorded ERPs during the presentation of each of four items (two numbers and two letters) during a simple comparison task (indicate the order of letters, in one condition, or numbers, in another). Data from only one electrode and one subject were presented, a very short epoch was used (510 msec), some data were discarded, no statistical analyses were performed after their PCA, and with only four elements there is a problem due to a "first item effect" (the initial item in any series, irrespective of condition, usually elicits a very large response).

Experiments designed to investigate P300 should meet several minimum

requirements, including the use of at least 3 electrodes (Fz, Cz, Pz) and a 1 second (or longer) recording epoch with appropriate amplification and filtering (see Donchin, Callaway, Cooper, Desmedt, Goff, & Hillyard, 1977; Duncan-Johnson & Donchin, 1979). The EOG must also be recorded and examined on each trial. Individual trials can then be rejected, or a correction procedure can be used (Gratton et al, 1983a). "Atypical" subjects or data should not be eliminated without justification, and when a PCA is performed, ANOVAs should be reported on the component scores.

Sanquist et al. (1980) presented two words to their subjects and then asked for a same-different judgment, based on similarity involving either orthography (are the two words both upper of lower case), phonology (do the words rhyme), or semantics (are the words synonyms). ERPs elicited by the second word in each pair were averaged into two classes based on the outcome of a later recognition test; i.e., words were divided into those subsequently recognized and not recognized. They report, however, that as they were forced to discard many trials due to artifacts only a small percentage of the data were analyzed. Therefore, separate averages were not calculated for "same" and "different" responses. The group composed of words subsequently recognized contained almost two times more "same" trials than "different". In the group composed of words that were subsequently unrecognized there were over four times more different trials than same. Since same responses produced a far larger P300 than different trials, the memory effect evident in their waveforms may be attributed to the discrepancy in the proportion of same and different trials within each memory group. It would appear, therefore, that there is at present no

substantive body of data that relates the ERP elicited by an item on its presentation to subsequent recall. Inasmuch as it is the original presentation of the stimulus, and its subsequent processing, that determines the strength and retrievablitiy of the item in memory, such data are likely to be of value. We report here a first attempt to determine if such is indeed the process.

In a series of experiments complementary to ours, Geiselman, Woodward, and Beatty (1982) combined psychophysiological and traditional measures to develop, and test, models of memory recall. In their second experiment they presented eight words simultaneously for 20 seconds and recorded eye movements, heart rate variability, and galvanic skin responses (GSRs). The later two measures, along with a single self report questionnaire, were used as indicators of a hypothetical construct they labeled "processing intensity." Processing intensity in their model accounted for 11% of the variance in recall from "long term store", independently of rehearsal strategy. (Processing intensity was not related to recall from the "short term store".) However, they averaged the two physiological indicators across trials, so were unable to examine, within subjects, the relationship between variations in these measures and recall. Ideally, one would want to record a measure of processing intensity _for each word_. The low temporal resolution of GSR and heart rate variability make these unsatisfactory for this purpose. ERPs may prove valuable, although simultaneous presentation of stimuli would present problems unless sophisticated eye monitoring equipment were used, and techniques perfected for recording from points in time identified by eye movements. This would permit concurrent assessment

of rehearsal strategies by examining fixation duration and fixation duration variability, as described by Geiselman et al. (1982).

GENERAL DISCUSSION

The P300 predicts recall better in group 1 than in the other two groups. When we examine the correspondence between the strategies subjects used to remember the words, performance, and P300, a coherent picture emerges. P300 provides information about the cognitive processing of an event that occurs only during the first second after its presentation, while processes that influence recall often continue for an extended period. The relationship between P300 and recall will thus depend on the nature of this extended mnemonic processing. Subjects who used simple rote strategies (group 1) recalled a low percentage of words, exhibited a strong von Restorff effect, and for these subjects the P300 elicited by a word predicted recall performance. On the other hand, subjects who used complex elaborative strategies (group 3) recalled a high percentage of words, showed little or no von Restorff effect, and P300 did not predict recall. In subjects who used elaborative strategies recall depends on the consequences of what Mandler (1979) calls interitem processing, which involves linking or associating several items together. The initial encoding of the word at input does not affect recall when retrieval depends on organizational processes that proceeded well beyond the coding of the surface attributes of the word. It is tempting to speculate that the slow wave may be associated with the beginning of these associative organizational processes. The slow wave began comparatively late (after 500 msec), predicted recall, and was greater in group 3 than in group 1. This is consonant with recent evidence

that slow waves appear in tasks requiring extended processing (Ruchkin & Sutton, 1983).

In subjects who used rote strategies, on the other hand, recall depends very much on the quality and nature of the encoding of initial attributes of the word. We suggest that the initial processing of input words is the same for all subjects, in part because the presentation of the isolates invoke the processing manifested by the P300 in all subjects (with no difference in the distribution of P300 amplitude between extreme groups). This processing is activated with differing intensities, reflected by varying P300 amplitudes across trials. These differing intensities reflect some critical recall-related attribute of the representation of the word in memory. If little further processing occurs P300 amplitude will be related to recall; the larger the P300 elicited by a word, the greater the probability that the word will later be recalled. However, if processing continues after the time frame reflected by the P300, and if this processing is beneficial for recall, then the relationship between P300 and recall may be obscured. There can be no doubt that extended processing <u>is</u> beneficial for recall, especially when this entails organizing or chunking the material. This idea is not new. William James (1890), in his chapter on memory, writes that, "all improvement of the memory lies in the line of <u>elaborating the associates</u> of each of several things to be remembered" (p. 663). He advocated "better remembering by better <u>thinking</u>" (p. 664).

It would be interesting to examine the <u>first</u> item in each list. These words are recalled more frequently than words in other positions (primacy effect), and generally elicit very large P300s ("first item effect"). We

should also be able to find a relation between P300 and recall for these items. Unfortunately, we were unable to perform this analysis because a large percentage of these trials was rejected due to muscle movements and other artifacts.

The results of this investigation are consistent with a two-phase model of the information processing system that leads to the subjects' recall. Phase 1 is driven by the processes that are invoked as the words are encoded and categorized. For some reason, the subroutine manifested by the P300 is activated by the words that are displayed in a deviant (isolated) font. Perhaps the need to activate a new set of feature detectors required for processing the font invokes the context-updating routine. This processing affects the representation of the word in long term memory, and seems to occur with equal frequency in all three groups of subjects, as there are no differences among groups in P300 amplitude, or in amplitude variability. We conceptualize representations in memory as multidimensional "traces" containing information on both semantic and nonsemantic attributes. Isolation in this study results in orthographic distinctiveness that is represented in one dimension of the memory trace; isolation may also increase the overall "activation" level. This first phase is identical for all subjects. The larger the P300 that a word elicits, the greater the activation in long term memory. The differences between the subjects appear in phase 2 - when subjects must try to recall the stimuli. Here the subjects' retrieval strategies play a crucial role. The subjects who rely on rote memorization seem to be aided in recalling a word by the fact the it has been isolated, due to their simple search strategies. These subjects

rely primarily on trace strength and on orthographic distinctiveness, and P300 is related to recall (see Hunt & Elliot, 1980, and their presentation of the "distinctiveness hypothesis"). On the other hand, the "elaborators" of group 3 do not rely on simple activation level for retrieval, or on the orthographic dimension of stimuli, and the amplitude of P300 for these subjects will thus be the same for recalled and for non recalled words.

Another way of explaining the group differences in the von Restorff effect is to think of the isolate as an organizational aid separating the list into two groups, an isolate and everything else. This was proposed originally by von Restorff herself (1933), and more recently by Bruce & Gaines (1976). Von Restorff wrote, in fact, that the magnitude of the isolation effect depends on the degree to which the organization of the sequence results in a separation between isolates and non-isolates. Von Restorff rarely used a single isolate in her experiments; Bruce and Gaines used four, and argue that a similar organizational hypothesis is sufficient to explain cases when only one isolate is presented. Isolation as an organizational aid will help rote memorizers, but not elaborators, who already are using a variety of effective organizational strategies. Other investigators present support for a selective attention-rehearsal hypothesis (Bellezza & Cheney, 1973; Cooper & Pantle, 1967; Rundus, 1971; Waugh, 1969). Isolates are more likely to be recalled, they argue, because they are held longer in mind and rehearsed more often. The evidence for this is quite shaky, however, as these investigators instruct subjects "to be sure to remember them" (the isolates)(Rundus, 1971, p.70). Other investigators have not found an increase in rehearsal for isolated items (Bruce & Gaines, 1976;

Einstein, Pellegrino, Mondani, & Battig, 1974; see also Hunt & Elliot, 1980, exp. 2).

Extensive individual differences are, or course, not unusual, when one looks for them. In a large study designed to examine the interrelationships among a variety of memory tasks, Underwood, Boruch, and Malmi (1978) encountered this problem: "The underlying individual differences in rate of associative learning appear to be so powerful that they dominate and obscure any relatively small amounts of variance due to individual differences on another factor, even if such variance exists"(p.415).

The relation between recall strategy and the von Restorff effect that we report here is consistent with other data that show that as the intrinsic organization of the list increases the von Restorff effect diminishes. When the list is constructed so that its items can be organized easily subjects will take advantage of this, overall recall will improve, and isolated items will no longer stand out. Bird (1980), for example, found a strong von Restorff effect only in "unrelated" lists, where each word was drawn from a different category, and not in "related" lists, where all words belonged to the same category (see also Bruce & Gaines, 1976). Similarly, Rosen, Richardson, & Saltz (1962) found a larger effect with words of low meaningfulness than with words of high meaningfulness, and Kothurkar (1956), who inserted numbers into a series of nonsense syllables or meaningful prose, found a far larger effect with the nonsense syllables. The ability to impose complex organizational schemes on a list depend both upon the nature of the list and the abilities of the individual. To the extent that such organization is imposed, the von Restorff effect will diminish.

Children engage in less associative encoding than adults, and there have been some reports that children show strong von Restorff effects (Cimbalo, Nowak, & Soderstrom, 1981). In incidental paradigms recall is not expected and less time is spent on organizational processes. One would thus expect the von Restorff effect to emerge under these conditions, and the relationship between P300 and recall should be clearer. The effect of incidental versus intentional designs on the von Restorff effect is not clear (Wallace, 1965), but we have found, in a preliminary study, that the relationship between memory and P300 may be clearer using incidental free recall (Karis et al., 1982). After several trials of a study-test recognition paradigm we unexpectedly asked for free recall. ERPs recorded to words presented during the study phase distinguished between words later recalled and words not later recalled only during the first free recall, which was unexpected, and not in subsequent free recalls.

The fact that there were no major differences between our three groups in recognition performance provides further support for the suggestion that all three groups interacted, at the initial encoding and storage phases, in the same manner with the stimuli. Recognition is likely to be affected more by the representation formed at the input phase than by mnemonic strategies. The recognition data are in accord with a study by Van Dam, Peek, Brinkerink, and Gorter (1974), who found a von Restorff effect in free recall but not in recognition. The strategies of subjects in group 1 that led to poor recall would not have led to poor recognition, for intraitem integration or organization of an event (which would occur during rote rehearsal) leads to familiarity and good recognition performance (Mandler,

1980, p. 255; Mandler, 1979; Tversky, 1973). In fact, the only significant difference among groups resulted from longer reaction times to new words recorded from subjects in group 3. In general, organization has a greater impact on free recall than recognition. Kintsch (1968), for example, found that the structure of his lists influenced recall but not recognition. Maintenance rehearsal, by facilitating associations between individual items and the list context, will thus improve recognition performance, but not recall (Crowder, 1976, p.387; Woodward, Bjcrk, & Jongeward, 1973; see also Geiselman & Bjork, 1980).

It may be useful to note that the precise pattern of our data, and in particular the individual differences, is not consistent with a class of accounts that attributes the effects of isolation, as well as the relationship between P300 and recall, to the hypothetical construct of "arousal" (see Roth, 1983). This view assumes that the larger the P300, the greater the arousal. Similar relations are presumed to hold between arousal and other psychophysiological measures (such as GSR or Heart Rate). Arousal, in turn, is supposed to assure better processing (except, of course, at extreme levels). The improved recall is correlated, according to this account, with increased P300 because both recall and P300 are "correlates" of arousal. This view suggests that no inferences regarding the process manifested by P300 can be made from our data. This argument is, however, weakened by the individual differences we have observed in recall, differences that did not correspond to differences between subjects in the initial amplitude, and the distribution of amplitudes, of the P300 elicited by the isolates. If P300 amplitude is an index . arousal then all subjects

were "aroused" to the same degree by the isolates. If degree of arousal is all that is needed to acount for differences in recall then it is difficult to see why different subjects show different strategy-dependent recall patterns. It seems neccessary to assume that changes in the amplitude of the P300 manifest processes that modulate the specific representations of words in memory. The modulations have a specific effect on recall probability, as a function of the subjects' mnemonic strategies.

The data presented in this paper serve to illustrate the ways in which the endogenous components of the ERP may be used, in conjunction with observations or overt responses, to enhance the analysis of human cognitive function. The manner in which the relation between P300 amplitude and recall varied with the subjects' strategies provides information that can be used to develop, and assess, more detailed models of storage and retrieval than can be derived from an analysis of the subject's reactions alone. It is in this fashion that ERPs can provide valuable information to the cognitive scientist.

evoked potentials. Canadian Journal of Psychology, 35, 201-212.

Cimbalo, R.S. (1978). Making something stand out: The isolation effect in memory performance. In M.M. Grunnenberg, P.E. Morris, and R.N. Sykes (Eds.), Practical aspects of memory . New York: Academic Press.

Cimbalo, R.S, Nowak, B.I., & Soderstrom, J.A. (1981). The isolation effect in children's short term memory. The Journal of General Psychology, 105, 215-223.

Coles, M.G.H., Gratton, G., Kramer, A.F., & Miller, G.A. (in press). Principles of signal acquisition and analysis. In M.G.H. Coles, E. Donchin, & S.W. Porges (Eds.), Psychophysiology: Systems, processes, and applications: Vol 1. Systems. New York: Guilford Press.

Cooper, E.H., & Pantle, A.J. (1967). The total-time hypothesis in verbal learning. Psychological Bulletin, 68, 221-224.

Crowder, R.G. (1976). Principles of learning and memory. Hillsdale: LEA.

Curry, S.H., Cooper, R., McCallum, W.C., Pocock, P.V., Papakostopoulos, D., Skidmore, S., & Newton, P. (1983). The principal components of auditory target detection. In A.W.K. Gaillard & W. Ritter (Eds.), Tutorials in event related potential research: Endogenous components (pp. 79-118). Amsterdam: North-Holland Publishing Company.

Detterman, D.K. (1975). The von Restorff effect and induced amnesia: Production by manipulation of sound intensity. Journal of Experimental Psychology: Human Learning and Memory, 1, 614-628.

Diaconis, P., & Efron, B. (1983). Computer intensive methods in statistics. Scientific American, 248(5), 116-130.

Donchin, E. (1969a). Data analysis techniques in average evoked potential

References

Adam, N., & Collins, G.I. (1978). Late components of the visual evoked potential to search in short-term memory. Electroencephalography and Clinical Neurophysiology, 44, 147-156.

Baddeley, A.D. (1981). The concept of working memory: A view of its current state and probable future development. Cognition, 10, 17-23.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.), The Psychology of Learning and Motivation, (Vol. 8). New York: Academic Press.

Bellezza, F., & Cheney, T.L. (1973). Isolation effect in immediate and delayed recall. Journal of Experimental Psychology, 99, 55-60.

Bird, C.P. (1980). The isolation effect as a function of unique processing orientation. Journal of Experimental Psychology: Human Learning and Memory, 6, 267-275.

Broadbent, D.E. (1981). Association lecture: From the percept to the cognitive structure. In J. Long and A. Baddeley (Eds.), Attention and Performance IX, Hillsdale, N J: Lawrence Erlbaum Associates.

Bruce, D., & Gaines, M.T. (1976). Tests of an organizational hypothesis of isolation effects in free recall. Journal of Verbal Learning and Verbal Behavior, 15, 59-72.

Chapman, R.M., McCrary, J.W., & Chapman, J.A. (1978). Short-term memory: The "storage" component of human brain responses predicts recall. Science, 202, 1211-1214.

Chapman, R.M., McCrary, J.W., & Chapman, J.A. (1981). Memory processes and

research. In E. Donchin & D.B. Lindsley (Eds.), Average evoked potentials: Methods, results, evaluations. NASA SP-19. Washington, D.C.: U.S. Government Printing Office, 199-236.

Donchin, E. (1969b). Discriminant analysis in average evoked response studies: The study of single trial data. Electroencephalography and Clinical Neurophysiology, 27, 311-314.

Donchin, E. (1979). Event-related brain potentials: A tool in the study of human information processing. In H. Begleiter (Ed.), Evoked Brain Potentials and Behavior . New York: Plenum Press, 13-75.

Donchin, E. (1981). Surprise!...Surprise?, Psychophysiology, 18, 493-513.

Donchin, E., Callaway, E., Cooper, R., Desmedt, J.E., Goff, W.R., Hillyard, S.A., & Sutton, S. (1977). Publication criteria for studies of evoked potentials (EP) in man. In J.E. Desmedt (Ed.) Attention, voluntary contraction and event-related cerebral potentials. Prog. clin. Neurophysiol., Vol. 1, 1-11.

Donchin, E., & Heffley, E. (1975). Minicomputers in the signal-averaging laboratory. American Psychologist, 38 , 449-461.

Donchin, E., & Heffley, E. (1978). Multivariate analysis of event-related potential data: A tutorial review. In D. Otto (Ed.), Multidisciplinary perspectives in event-related potential research, EPA-600/9-7.-043, Washington D.C.: U.S. Government Printing Office, 555-572.

Donchin, E., Ritter, W., & McCallum, C. (1978). Cognitive Psychophysiology: the endogenous components of the ERP. In E. Callaway, P. Tueting, and S. Koslow (Eds.), Brain Event-Related Potentials in Man . New York: Academic Press, 349-441.

Duncan-Johnson, C.C., & Donchin, E. (1977). On quantifying surprise: The variation in event-related potentials with subjective probability, Psychophysiology, 14 , 456-467.

Duncan-Johnson, C.C., & Donchin, E. (1979). The time constant in P300 recording. Psychophysiology, 16, 53-55.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7, 1-26.

Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician, 37, 36-48.

Einstein, G.O., Pellegrino, J.W., Mondani, M.S., & Battig, W.F. (1974). Free recall performance as a function of overt rehearsal frequency. Journal of Experimental Psychology, 103, 440-449.

Ericsson, K.A., & Simon, H.A. (1980). Verbal reports as data. Psychological Review, 87, 215-251.

Fitzgerald, P.G., & Picton, T.W. (1981). Temporal and sequential probability in evoked potential studies. Canadian Journal of Psychology, 35, 188-200.

Ford, J.M., Roth, W.T., Mohs, R.C., Hopkins, W.F. III, & Kopell, B.S. (1979). Event-related potentials recorded from young and old adults during a memory retrieval task. Electroencephalography and Clinical Neurophysiology, 47, 450-459.

Geiselman, R.E., & Bjork, R.A. (1980). Primary versus secondary rehearsal in imagined voices: Differential effects on recognition. Cognitive Psychology, 12, 188-205.

Geiselman, R.E., Woodward, J.A., & Beatty, J. (1982). Individual differences in verbal memory performance: A test of alternative information-processing models. Journal of Experimental Psychology: General, 111, 109-134.

Gomer, F.E., Spicuzza, R.J., & O'Donell, R.D. (1976). Evoked potential correlates of visual item recognition during memory-scanning tasks. Physiological Psychology, 4, 61-65.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983a). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55, 468-484.

Gratton, G., Coles, M.G.H., & Donchin, E. (1983b, September). Filtering for spatial distribution: A new approach (Vector Filter). Paper presented at the meeting of the Society for Psychophysiological Research, Asilomar, CA.

Grubin, M.L., Bauer, J.A., Jr., Walker, E.C.T. (1976). Alice User's Guide, Alice Associates, 29 Wellesley Avenue, Natick, MASS, 01760.

Heffley, E.F. (1981, April) Event-related brain potentials in visual monitoring tasks: Selective attention, display format, and event frequency. Paper presented at the Eastern Psychological Association Annual Meeting, New York City, NY.

Hillyard, S.A., Squires, K.C., Bauer, J.W., & Lindsay, P.H. (1971). Evoked potential correlates of auditory signal detection. Science, 172, 1357-1360.

Horst, R.L., & Donchin, E. (1980). Beyond averaging. II. Single-trial classification of exogenous event-related potentials using stepwise

discriminant analysis. Electroencephalography and Clinical Neurophysiology, 48, 113-126.

Hunt, E. (1980, October). The validity of assumptions underlying the use of PCA and ANOVA in ERP research. Paper resented at the meeting of the Society for Psychophysiological Research, Vancouver, Canada.

Hunt, R.R., & Eliot, J.M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. Journal of Experimental Psychology: General, 109, 49-74.

James, W. (1950; original copyright, 1890). The principles of psychology (Authorized Edition), Volume One. N.Y.: Dover Publications, Inc.

Jasper, H.H. (1958). Report of the committee on methods of clinical examination in electroencephalography, Electroencephalography and Clinical Neurophysiology, 10 , 370.

Johnson, R.E. Jr., Donchin, E. (1978). On how P300 amplitude varies with the utility of the eliciting stimuli. Electroencephalography and Clinical Neurophysiology, 44, 424-437.

Karis, D., Bashore, T., Fabiani, M., & Donchin, E. (1982). P300 and Memory. Psychophysiology, 19, 328, (Abstract).

Kellogg, R.T. (1982). When can we introspect accurately about mental processes? Memory and Cognition, 10, 141-144.

Kintsch, W. (1968). Recognition and free recall of organized lists. Journal of Experimental Psychology, 78, 481-487.

Koffka, K. (1935). Principles of Gestalt Psychology. New York: Harcourt, Brace, and Company.

Kohler, W. (1940). Dynamics in Psychology. New York: Liveright Publishing

Corporation.

Kothurkar, V.K. (1956). Learning and retention of an isolated number on the background of meaningful material. Indian Journal of Psychology, 31, 59-62.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time, Science, 197 , 792-795.

Mandler, G. (1979). Organization and repetition: Organizational principles with special reference to rote learning. In L.G. Nilsson (Ed.), Perspectives on Memory Research. Hillsdale, N.J.: Erlbaum.

Mandler, G. (1980). Recognizing: The judgment of pervious occurrence. Psychological Review, 87, 352-271.

Marsh, G.R. (1975). Age differences in evoked potential correlates of a memory scanning process. Experimental Aging Research, 1, 3-16.

Morris, P.E. (1981). Why Evans is wrong in criticizing introspective reports of subject strategies. British Journal of Psychology, 72, 465-468.

Naatanen, R., & Gaillard, A.W.K. (1983). The orienting reflex and the N2 deflection of the event-related potential (ERP). In A.W.K. Gaillard & W. Ritter (Eds.), Tutorials in event related potential research: Endogenous components (pp. 119-141). Amsterdam: North-Holland Publishing Company.

Nageishi, Y., & Shimokochi, M. (1980). Event related potentials varied with the contents of short-term memory. Osaka University, Journal of Human Science, 6, 81-99.

Neisser, U. (1976). Cognition and Reality. San Francisco: W.H. Freeman and Company.

Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.

Parasuraman, R., & Beatty, J. (1980). Brain events underlying detection and recognition of weak sensory signals. Science, 210, 80-83.

Parasuraman, R., Richer, F., & Beatty, J. (1982). Detection and recognition: Concurrent processes in perception. Perception and Psychophysics, 31, 1-12.

Paul, D.D., & Sutton, S. (1972). Evoked potential correlates of response criterion in auditory signal detection. Science, 177, 362-364.

Pribram, K.H., & McGuinness, D. (1975). Arousal, activation, and effort in the control of attention. Psychological Review, 82, 116-149.

Restorff, H. von (1933). Uber die Wirkung von Bereichsbildungen im Spurenfeld, Psychologische Forschung, 18 , 299-342.

Rosen, H., Richardson, D.H., & Saltz, E. (1962). Supplementary report: Meaningfulness as a differentiation variable in the von Restorff effect. Journal of Experimental Psychology, 64 , 327-328.

Roth, W.T. (1983). A comparison of P300 and skin conductance response. In A.W.K. Gaillard & W. Ritter (Eds.), Tutorials in event related potential research: Endogenous components (pp. 177-199). Amsterdam: North-Holland Publishing Company.

Roth, W.T., Kopell, B.S., Tinklenberg, J.R., Darley, C.F., Sikora, R., & Vesecky, T.B. (1975). The contingent negative variation during a memory retrieval task. Electroencephalography and Clinical Neurophysiology,

38, 171-174.

Roth, W.T., Tinklenberg, J.R. & Kopell, B.S. (1977). Ethanol and marihuana effects on event-related potentials in a memory retrieval paradigm. Electroencephalography and Clinical Neurophysiology, 42, 381-388.

Roth, W.T., Rothbart, R.M., & Kopell, B.S. (1978). The timing of CNV resolution in a memory retrieval task. Biological Psychology, 6, 39-49.

Ruchkin, D.S., & Sutton, S. (1983). Positive slow wave and P300: Association and disassociaion. In A.W.K. Gaillard & W. Ritter (Eds.), Tutorials in event related potential research: Endogenous components (pp. 233-250). Amsterdam: North-Holland Publishing Company.

Rundus, D. (1971). Analysis of rehearsal processes in free recall. Journal of Experimental Psychology, 89, 63-77.

Sanquist, T.F., Rohrbaugh, J.W., Syndulko, K., & Lindsley, D.B. (1980). Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. Psychophysiology, 17, 568-576.

Smith, E.R., & Miller, F.D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. Psychological Review, 85, 355-362.

Sokolov, E.N. (1963). Perception and the conditioned reflex. New York: Macmillan.

Sokolov, E.N. (1969). The modeling properties of the nervous system. In I. Maltzman and M. Cole (Eds.), Handbook of Contemporary Soviet Psychology. New York: Basic Books, 671-704.

Sokolov, E.N. (1975). The neuronal mechanisms of the Orienting Reflex. In E.N. Sokolov and O.S. Vimogradova (Eds.), Neuronal Mechanisms of the

Orienting Reflex . Hillsdale: LEA, 217-235.

Squires, K.C., Squires, N.K., & Hillyard, S. (1975a). Decision-related cortical potentials during an auditory signal detection task with cued observation intervals. Journal of Experimental Psychology: Human Perception and Performance, 1(3), 268-279.

Squires, K.C., Squires, N.K., & Hillyard, S. (1975b). Vertex evoked potentials in a rating-scale detection task: Relation to signal probability. Behavioral Biology, 13, 21-34.

Squires, N.K., Squires, K.C., & Hillyard, S.A. (1975c). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. Electroencephalography and Clinical Neurophysiology, 38, 387-401.

Squires, K.C., Wickens, C., Squires, N.K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. Science, 193 , 1142-1146.

Stanny, R.R., & Elfner, L.F. (1980). A short-term memory influence on the N1 response of cerebral cortex. Journal of Experimental Psychology: Human Perception and Performance, 6, 321-329.

Sutton, S., Braren, M., Zubin, J., & John, E.R. (1965). Evoked-potential correlates of stimulus uncertainty. Science, 150, 1187-1188.

Toglia, M.P., & Battig, W.F. (1978). Handbook of Semantic Word Norms, Hillsdale: LEA.

Tulving, E. (1968). When is recall higher than recognition. Psychonomic Science, 10, 53-54.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. Psychological Review, 80, 352-373.

Tversky, B. (1973). Encoding processes in recognition and recall. Cognitive Psychology, 5, 275-287.

Underwood, B.J., Boruch, R.F., & Malmi, R.A. (1978). Composition of episodic memory. Journal of Experimental Psychology: General, 107, 393-419.

Van Dam, G., Peeck, J., Brinkerink, M., & Gorter, U. (1974). The isolation effect in the free recall and recognition. American Journal of Psychology, 87 , 497-504.

Wagner, A.R. (1976). Priming in STM: An information processing mechanism for self-generated or retrieval-generated depression in performance. In T.J. Tighe, and R.N. Leaton (Eds.), Habituation perspectives from child development, animal behavior, and neurophysiology . Hillsdale: LEA.

Wallace, W.P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon, Psychological Bulletin, 63 , 410-424.

Warren, L.R. (1980). Evoked potential correlates of recognition memory. Biological Psychology, 11, 21-35.

Wastell, D.G. (1981). On the correlated nature of evoked brain activity: Biophysical and statistical considerations. Biological Psychology, 13, 51-69.

Watkins, M.J., & Todres, A.K. (1978). On the relation between recall and recognition. Journal of Verbal Learning and Verbal Behavior, 17, 621-633.

Watkins, M.J., & Tulving, E. (1975). Episodic memory: When recognition fails. Journal of Experimental Psychology: General, 104, 5-29.

Waugh, N.C. (1969). Free recall of conspicuous items. _Journal of Verbal Learning and Verbal Behavior_, _8_, 448-456.

White, P. (1980). Limitations on verbal reports of internal events: A refutation of Nisbett and Wilson and of Bem. _Psychological Review_, _87_, 105-112.

Woodward, A.E., Bjork, R.A., & Jongeward, R.H. (1973). Recall and recognition as a function of primary rehearsal. _Journal of Verbal Learning and Verbal Behavior_, _12_, 608-617.

Author Notes

A partial report on this study was presented at the Twenty-second Annual Meeting of the Society for Psychophysiological Research, Minneapolis, October 21-24, 1982.

This research was supported by the Office of Naval Research, contract #N000-14-76-C-0002, to E. Donchin, with funds provided by the Defense Advanced Research Projects Agency, and by the Air Force Office of Scientific Research, contract #F49620-79-C-0233. We thank Drs. Craig Fields and Al Fregly for their support. The first author was supported by an NIMH training grant, MH15128. We are indebted to Michael Gabriel, Earl Hunt, Art Kramer, David Strayer, and three anonymous reviewers for helpful comments on an earlier draft. Gabriele Gratton provided helpful statistical advice.

Address requests for reprints to: Emanuel Donchin, Psychology Department, University of Illinois, 603 East Daniel, Champaign, Illinois, 61820.

Footnotes

1. The "P" in P300 refers to polarity (positive). The "300" refers to the latency of the peak in msec (measured from stimulus presentation), because this is when it was first observed (Sutton, Braren, Zubin, & John, 1965). However, since P300 is elicited after stimulus evaluation is completed (Kutas, McCarthy, & Donchin, 1977) its latency varies widely, and often exceeds 500 msec in complex paradigms. Scalp distribution is important in identifying P300; it is usually largest parietally (Pz), slightly smaller centrally (Cz), and very small frontally (Fz).

2. "Working memory refers to the role of temporary storage in information processing" (Baddeley, 1981, p. 17; see also Baddeley & Hitch, 1974). The concept of working memory emphasizes function, while short term memory has often been used to refer to a hypothetical structure. We prefer working memory, and will use it throughout this paper.

3. For example, with two equiprobable events, A and B, the P300 elicited by the last A in a sequence will be larger in the sequence BA than AA. Similarly, in third order series, P300 amplitude will decrease from BBA to ABA to BAA to AAA.

4. Neisser's (1976) discussion of the "perceptual cycle" is very similar to this formulation.

5. A series of events that can be divided into discrete classes is called an "oddball" when one event (or class of events) is much rarer than the other (although sometimes even series with 50-50 probability are called oddballs). Since such paradigms have been used extensively, typical ERPs can be easily recognized, and subjects producing anomalous waveforms spotted. In addition, trials from an oddball can be used to create a discriminant function that is then used to assign a P300 amplitude score to individual trials in experimental conditions. In the Results section we describe the use of such a discriminant analysis.

6. In the experimental lists isolated words were either larger or smaller than the other words. There was no difference in recall between these two types of isolates.

7. Performance and the von Restorff index are not independent, but the expected value for the correlation _is_ zero. After dividing performance into its constituent parts (recall of isolates, recall of non-isolates in position 6-10, recall of non-isolates from other positions), it can be demonstrated that the covariance between P and VRI is equal to zero.

8. Some researchers report a decrement in recall for items on either side of an isolate (Detterman, 1975), although often no effect is found (Cimbalo, 1978; Wallace, 1965). We examined the recall of the two items before and after an isolated word with words from the control lists. To control for serial position effects this analysis was performed only on words in position 7 through 9, because these positions were common to words two before the isolate (positions 4 through 9) and two after (positions 7

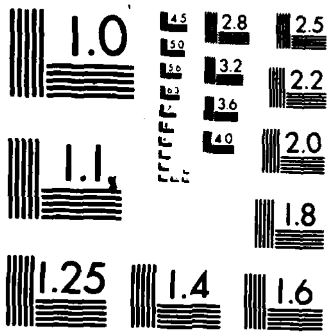
through 12). There was no difference between the recall of these words and control words in the same positions.

9. Geiselman, Woodward, & Beatty (1982), for example, found correlations of .74 and .82 between strategies, based on verbal reports, and two free recall performance measures.

10. We label this component a "frontal positive slow wave" to distinguish it from the more typical slow wave distribution reported in the past (negative frontally, becoming more and more positive as one moves back across the scalp; see Squires, Squires, & Hillyard, 1975c, and Ruchkin & Sutton, 1983).

AD-A137 779 THE EVENT RELATED BRAIN POTENTIAL AS AN INDEX OF INFORMATION PROCESSING C..(U) ILLINOIS UNIV CHAMPAIGN COGNITIVE PSYCHOPHYSIOLOGY LAB E DONCHIN ET AL. OCT 83 2/2

UNCLASSIFIED CPL83-4 AFOSR-TR-84-0051 F49620-79-C-0233 F/G 5/10 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Table 1

Individual and Group Performance in the Free and Grand Recall (Percentages)

	Free Recall			Grand Recall		
S#	von Restorff Index*[a]	Perfor-mance*[b]	Improve-ment*[c]	von Restorff Index*	Perfor-mance*	Improve-ment
Group 1						
2	33	46	0	8	10	-2
11	30	30	-5	12	8	-3
6	29	42	2	14	9	1
M(SD)	31(2)	39(8)	-1(4)	11(3)	9(1)	-1(2)
Group 2						
12	19	56	3	4	14	3
10	17	44	5	12	11	3
8	16	49	5	3	8	1
7	14	54	13	2	12	3
9	13	50	8	5	8	-1
4	11	47	9	0	9	1
M(SD)	15(3)	50(4)	7(4)	4(4)	10(2)	2(2)

Group 3							
3	5	61	10		0	12	2
1	-3	63	11		4	13	9
5	-6	65	5		-8	19	-1
M(SD)	-1(6)	63(2)	9(3)		-1(6)	15(4)	3(5)
All Subjects							
M(SD)	15(12)	51(10)	5(5)		5(6)	11(3)	1(3)

[a] von Restorff Index: % isolates recalled minus % non-isolates recalled (position 6-10)

[b] Performance: % recalled from all positions

[c] Improvement: % performance in session 2 - % performance in session 1

* Groups differ significantly on this variable ($p < .05$)

Table 2

Correlations Between Recall Measures

Variables	Free Recall			Grand Recall		
	VRI[a]	Perfor-mance	Improve-ment	VRI	Perfor-mance	Improv-ment
Free Recall						
VRI	1.00					
Performance	-0.84*	1.00				
Improvement	-0.70*	0.68*	1.00			
Grand Recall						
VRI	0.79*	-0.78*	-0.53	1.00		
Performance	-0.65*	0.76*	0.23	-0.65*	1.00	
Improvement	-0.50	0.53	0.64*	-0.08	0.27	1.00

Note. N = 12

[a] VRI = von Restorff Index

* $p < .05$

Table 3

Reaction Times and Error Rates from the Recognition Test

	REACTION TIMES[a] (msec)			ERROR RATE[b](%)		
	Isolates	Non-Isolates	New	Isol.	Non-Isolates	New
Group 1	725	814	862	29	33	27
	(36)	(20)	(27)	(6)	(8)	(18)
Group 2	740	764	847	29	33	21
	(128)	(157)	(102)	(13)	(10)	(11)
Group 3	799	823	1040	22	18	32
	(56)	(66)	(52)	(7)	(5)	(16)
Total	751	791	899	27	29	25
	(100)	(119)	(113)	(11)	(11)	(15)

[a] Group means are presented, calculated from individual medians. Standard deviations are in parentheses.

[b] Errors for isolates and non-isolates can be considered misses, errors for new words false alarms.

Table 4

Analysis of Variance on the PCA Component Scores[a]

SOURCE	DF1/ DF2	MEAN SQUARE	F	P VALUE
GR (GROUP)	2/ 9	6.0091	0.6773	0.5321
	2/ 9	25.9304	5.7074	0.0251
WO (WORD)	2/18	14.1962	20.7290	0.0000
	2/18	3.2566	3.8333	0.0410
GR*WO	4/18	0.7096	1.0362	0.4158
	4/18	0.4895	0.5761	0.6836
ME (MEMORY)	1/ 9	2.1971	14.0859	0.0045
	1/ 9	19.7069	10.7835	0.0095
GR*ME	2/ 9	0.9010	5.7761	0.0243
	2/ 9	1.1892	0.6507	0.5446
WO*ME	2/18	0.0470	0.5746	0.5729
	2/18	0.3830	0.7439	0.4893
GR*WO*ME	4/18	0.0744	0.9089	0.4798
	4/18	0.1347	0.2616	0.8987
EL (ELECTRODE)	2/18	13.7051	37.1394	0.0000
	2/18	2.6610	2.2844	0.1306
GR*EL	4/18	1.0478	2.8394	0.0549
	4/18	2.1330	1.8312	0.1668
WO*EL	4/36	2.8820	42.8976	0.0000
	4/36	1.0860	7.3942	0.0002
GR*WO*EL	8/36	0.1267	1.8854	0.0930
	8/36	0.1864	1.2688	0.2900

ME*EL	2/18	0.0036	0.1984	0.8218
	2/18	1.0476	6.6450	0.0069
GR*ME*EL	4/18	0.0099	0.5548	0.6981
	4/18	0.1058	0.6714	0.6203
WO*ME*EL	4/36	0.0023	0.1568	0.9587
	4/36	0.1249	0.8066	0.5291
GR*WO*ME*EL	8/36	0.0567	3.9264	0.0020
	8/36	0.3666	2.3681	0.0368
GR*SS	9	8.8726		
	9	4.5433		
GR*SS*WO	18	0.6848		
	18	0.8496		
GR*SS*ME	9	0.1560		
	9	1.8275		
GR*SS*WO*ME	18	0.0818		
	18	0.5149		
GR*SS*EL	18	0.3690		
	18	1.1648		
GR*SS*WO*EL	36	0.0672		
	36	0.1469		
GR*SS*ME*EL	18	0.0179		
	18	0.1576		
GR*SS*WO*ME*EL	36	0.0144		
	36	0.1548		

[a] Each entry includes results for component 1 (P300 - first line) followed by the results for component 2 (frontal positive slow wave - second line).

Figure Captions

Figure 1. Experimental design.

Figure 2. The von Restorff index is plotted against overall performance for each subject. Grand means for each index are also shown ($\overline{VRI}$, $\overline{P}$). Subjects were divided into three groups on the basis of their von Restorff index.

Figure 3. Free recall serial position curves for each group, with the isolates plotted separately. Control words are from lists that contained no isolated item.

Figure 4. The von Restorff index is plotted against improvement in free recall (session 2 - session 1) for each subject. Grand means for each index are also shown ($\overline{VRI}$ and $\overline{I}$).

Figure 5. Individual averages at Pz for isolates, non-isolates, and controls. Circled numbers refer to individual subjects and correspond to the numbers used in Figures 2 and 4.

Figure 6. Group averages at Pz for the three classes of words. All words were presented in position 6 through 10. Each average is divided into recalled vs. not recalled.

Figure 7. Group averages for isolates at the three electrode sites. Each average is divided into recalled vs. not recalled.

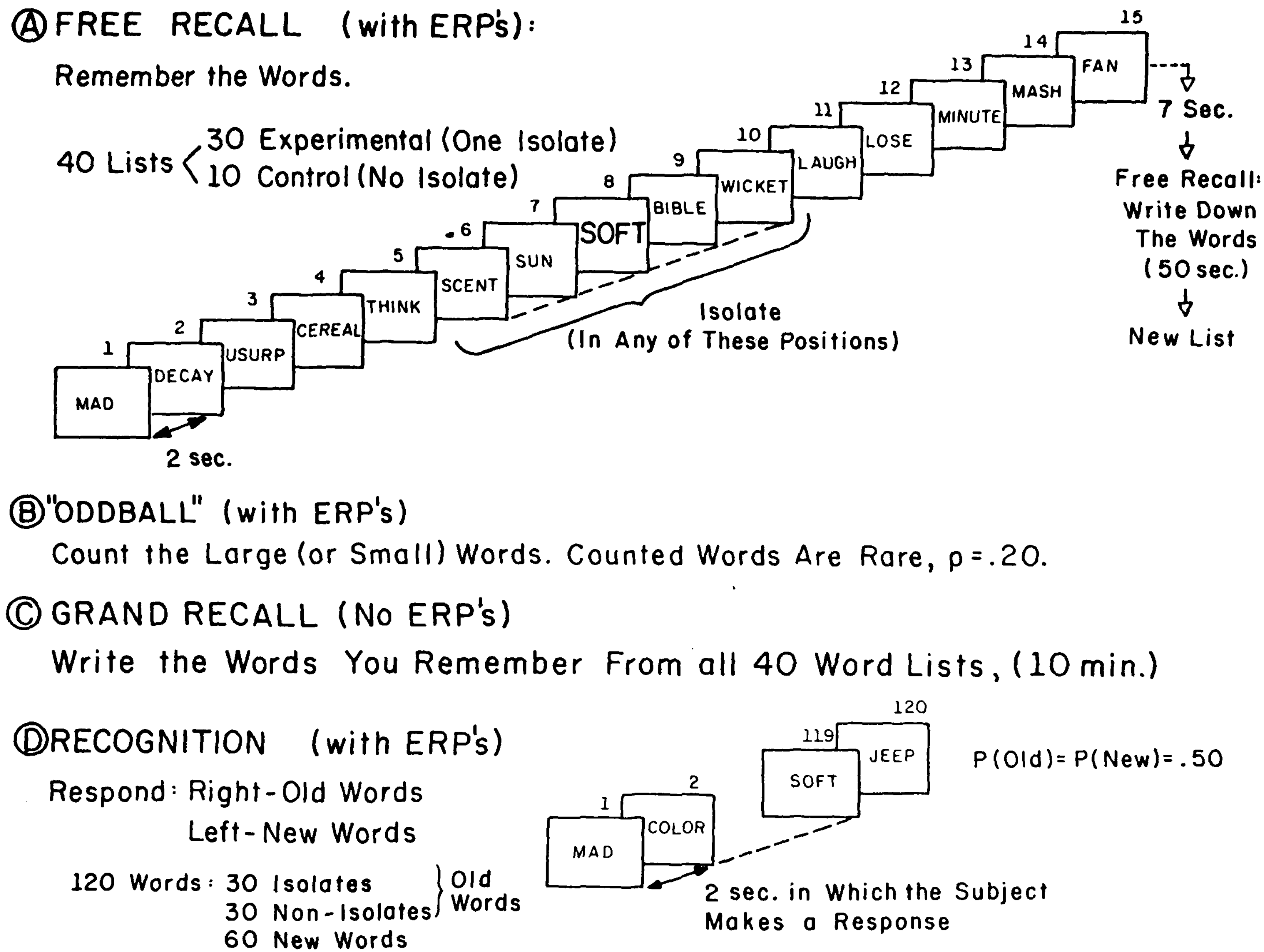
Figure 8. Component loadings for four components derived from a Principal Components Analysis (using the covariance matrix and varimax rotation) of the average waveform data elicited by words in position 6 through 10.

Figure 9. Component scores for isolated words for the first two components.

Figure 10. Grand average waveforms for 12 subjects for the three classes of words at the three electrode sites. This is a reaveraging of the free recall data based on recall in both the free recall and grand recall phases.

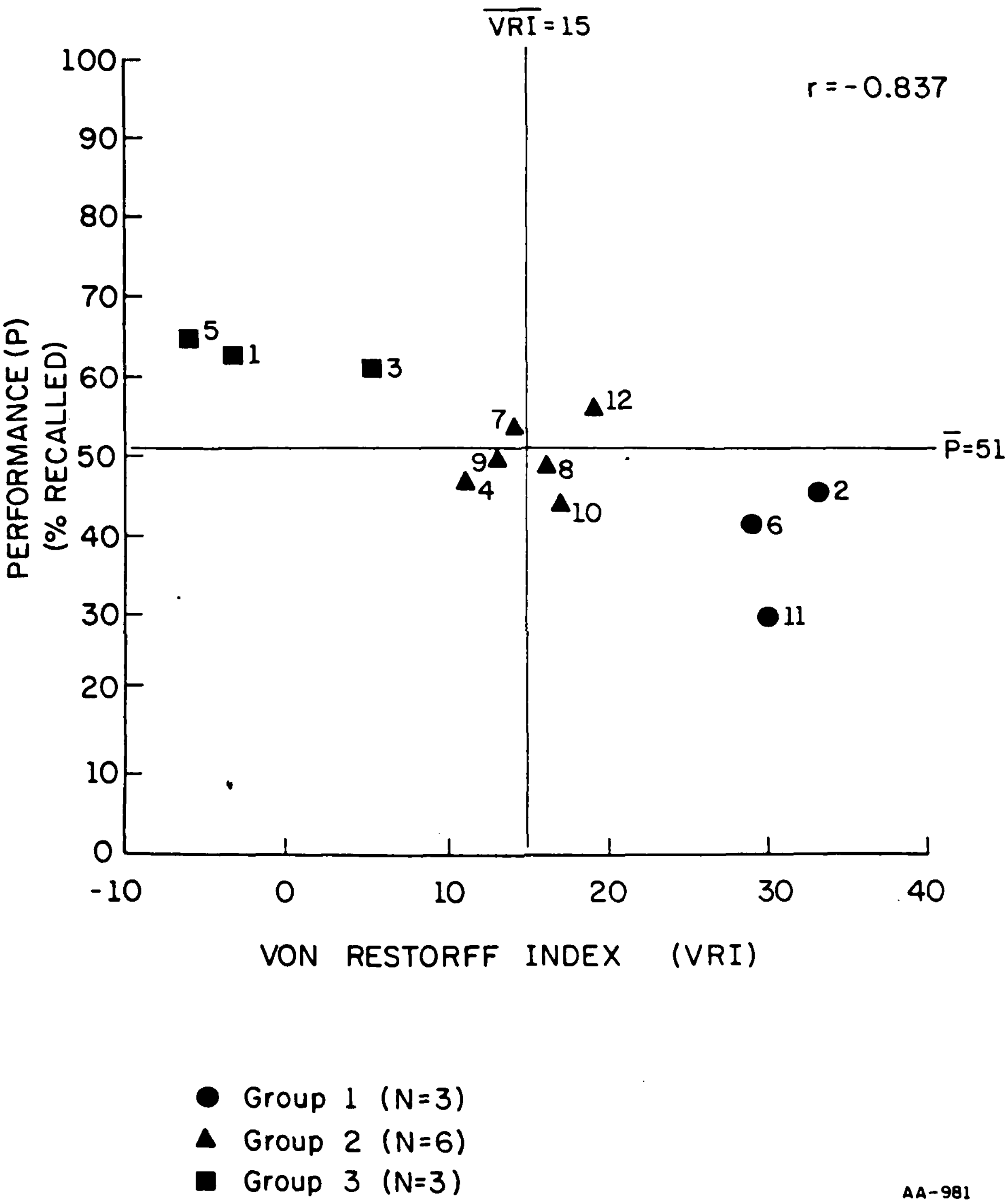
Figure 11. Grand averages for all subjects for the isolates at Pz. This figure depicts further reaveraging on the free recall data on the basis of all three performance measures: free recall, grand recall, and recognition.

Figure 12. Grand average waveforms for 12 subjects elicited by words presented during the recognition test. Averages are presented for three classes of words at the three electrode sites, and are divided into those correctly and incorrectly recognized.



Ⓐ FREE RECALL (with ERP's):
Remember the Words.
40 Lists
30 Experimental (One Isolate)
10 Control (No Isolate)
1
MAD
2
DECAY
3
USURP
4
CEREAL
5
THINK
6
SCENT
7
SUN
8
SOFT
9
BIBLE
10
WICKET
11
LAUGH
12
LOSE
13
MINUTE
14
MASH
15
FAN
2 sec.
Isolate
(In Any of These Positions)
7 Sec.
Free Recall:
Write Down
The Words
(50 sec.)
New List
Ⓑ "ODDBALL" (with ERP's)
Count the Large (or Small) Words. Counted Words Are Rare, p = .20.
Ⓒ GRAND RECALL (No ERP's)
Write the Words You Remember From all 40 Word Lists, (10 min.)
Ⓓ RECOGNITION (with ERP's)
Respond: Right-Old Words
Left-New Words
120 Words: 30 Isolates
30 Non-Isolates
Old Words
60 New Words
1
MAD
2
COLOR
119
SOFT
120
JEEP
P(Old) = P(New) = .50
2 sec. in Which the Subject
Makes a Response

Figure 1



$\overline{VRI}=15$
r = -0.837
$\bar{P}=51$
PERFORMANCE (P) (% RECALLED)
100
90
80
70
60
50
40
30
20
10
0
-10
0
10
20
30
40
VON RESTORFF INDEX (VRI)
5
1
3
7
12
9
8
4
10
2
6
11
Group 1 (N=3)
Group 2 (N=6)
Group 3 (N=3)
AA-981

Figure 2

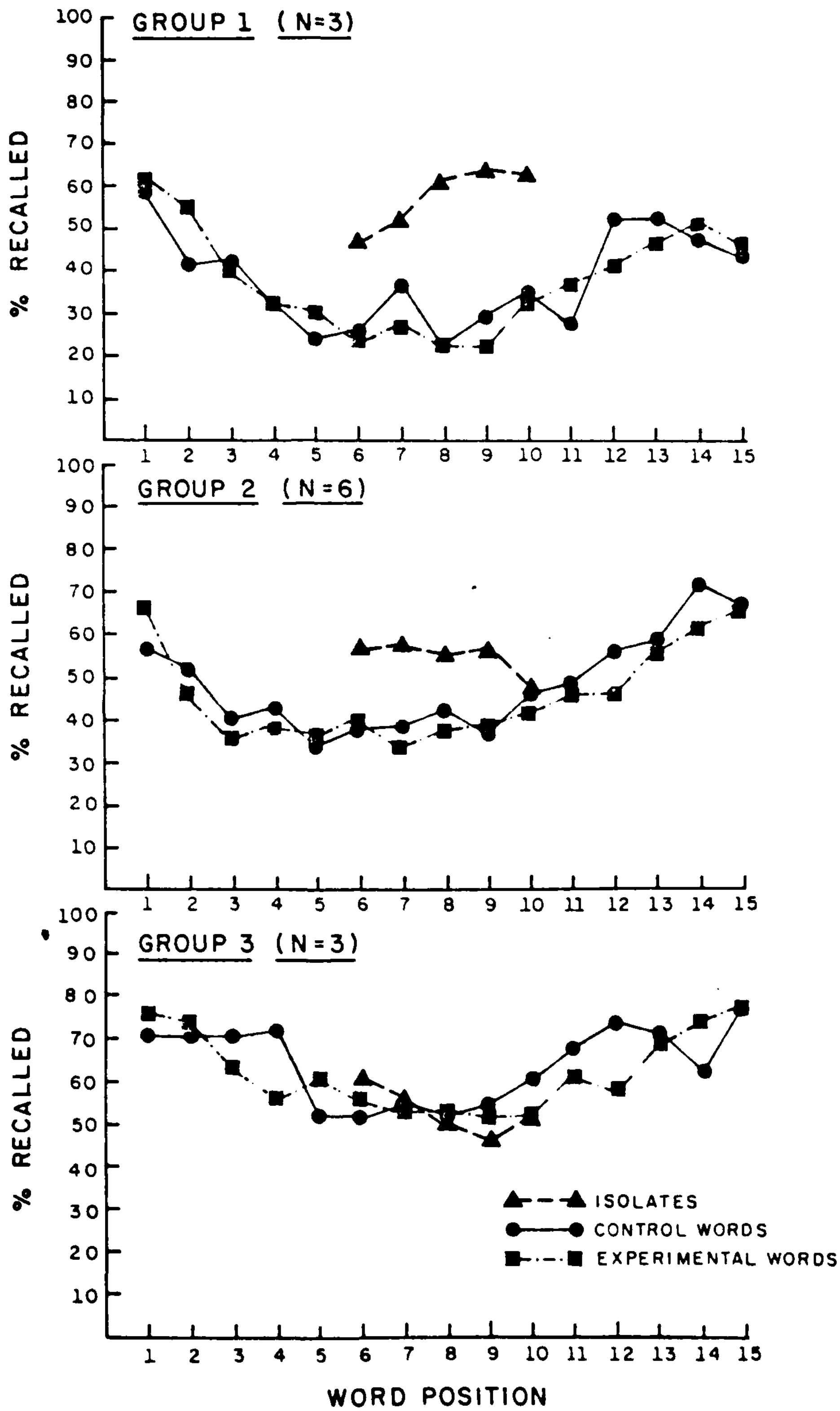

GROUP 1 (N=3)
GROUP 2 (N=6)
GROUP 3 (N=3)
% RECALLED
WORD POSITION
ISOLATES
CONTROL WORDS
EXPERIMENTAL WORDS

Figure 3

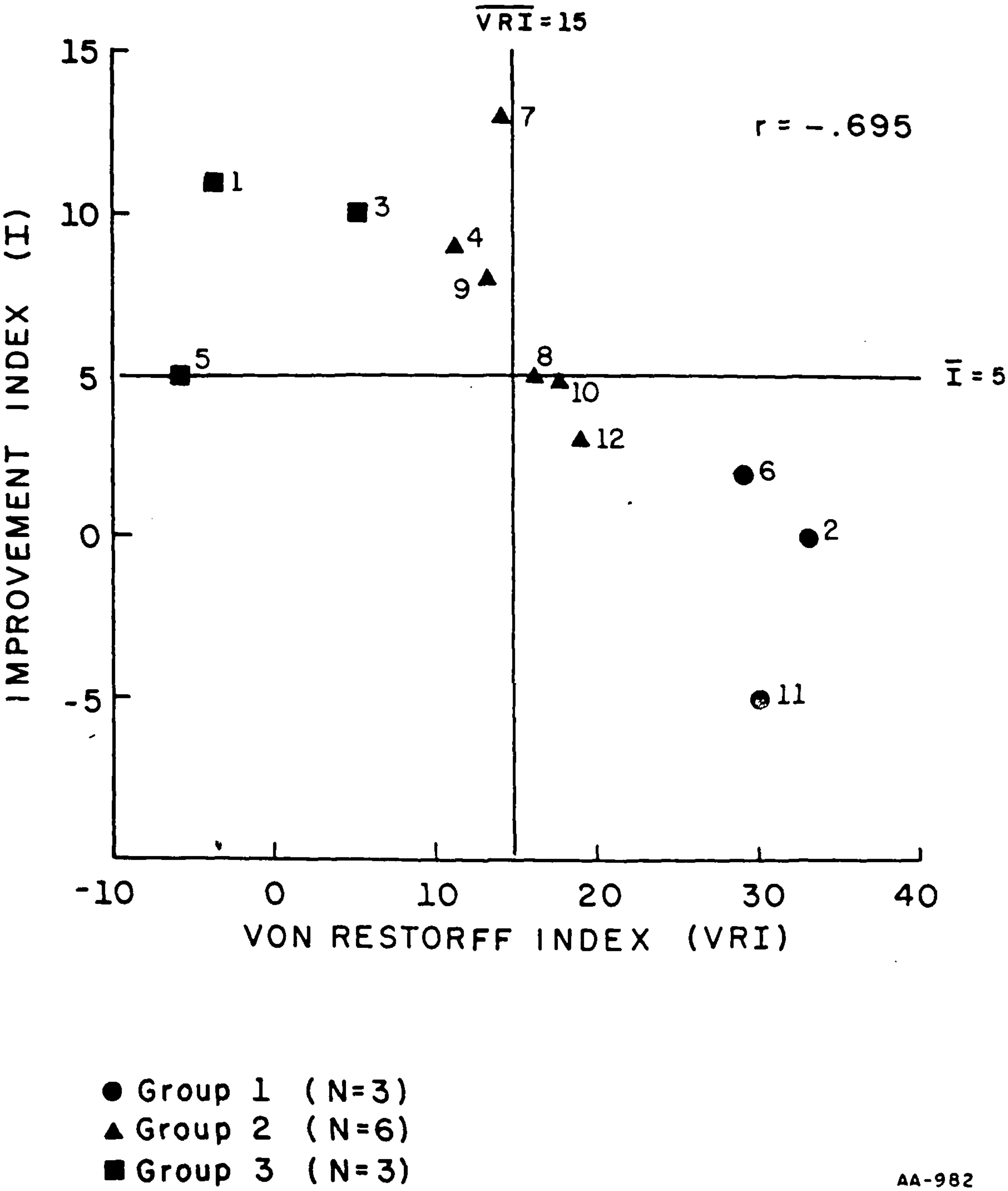
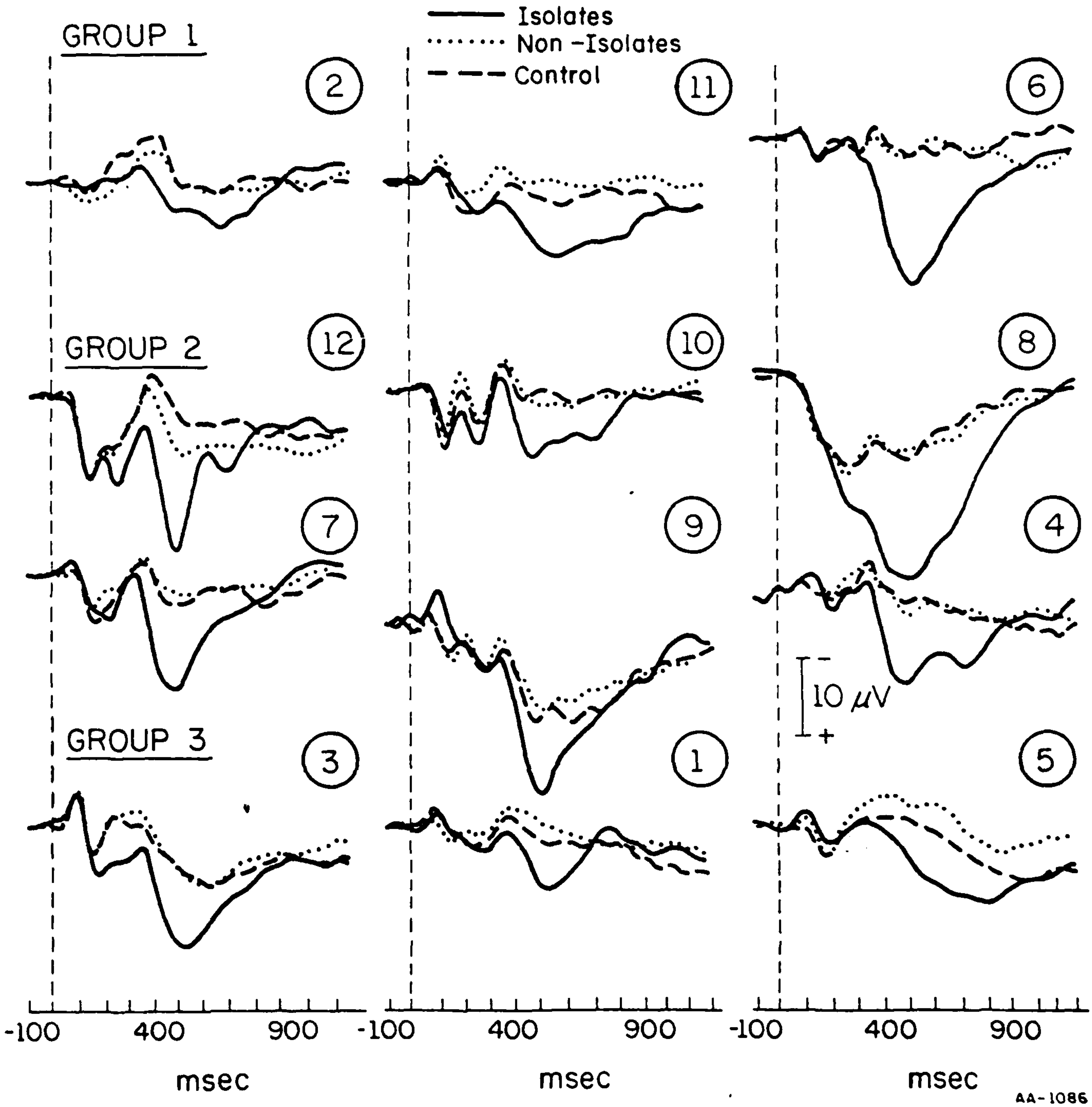

VRI = 15
r = -.695
I = 5
IMPROVEMENT INDEX (I)
VON RESTORFF INDEX (VRI)
15
10
5
0
-5
-10
0
10
20
30
40
1
2
3
4
5
6
7
8
9
10
11
12
● Group 1 (N=3)
▲ Group 2 (N=6)
■ Group 3 (N=3)
AA-982

Figure 4



GROUP 1
GROUP 2
GROUP 3
Isolates
Non -Isolates
Control
2
11
6
12
10
8
7
9
4
3
1
5
10 μV
-100
400
900
msec
AA-1086

Figure 5

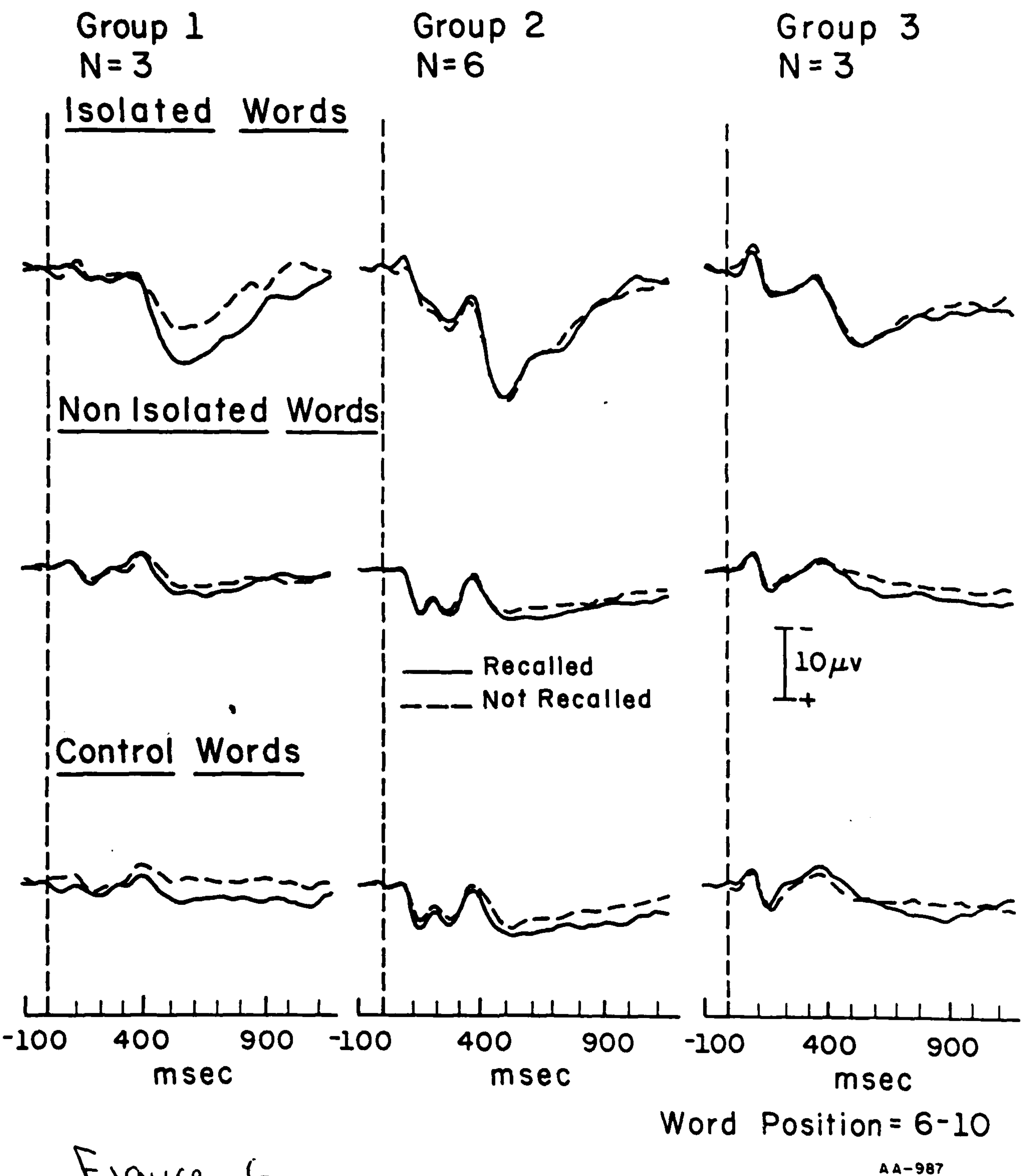

Group 1
N=3
Group 2
N=6
Group 3
N=3
Isolated Words
Non Isolated Words
Control Words
Recalled
Not Recalled
10μv
-100
400
900
msec
Word Position = 6-10
AA-987

Figure 6

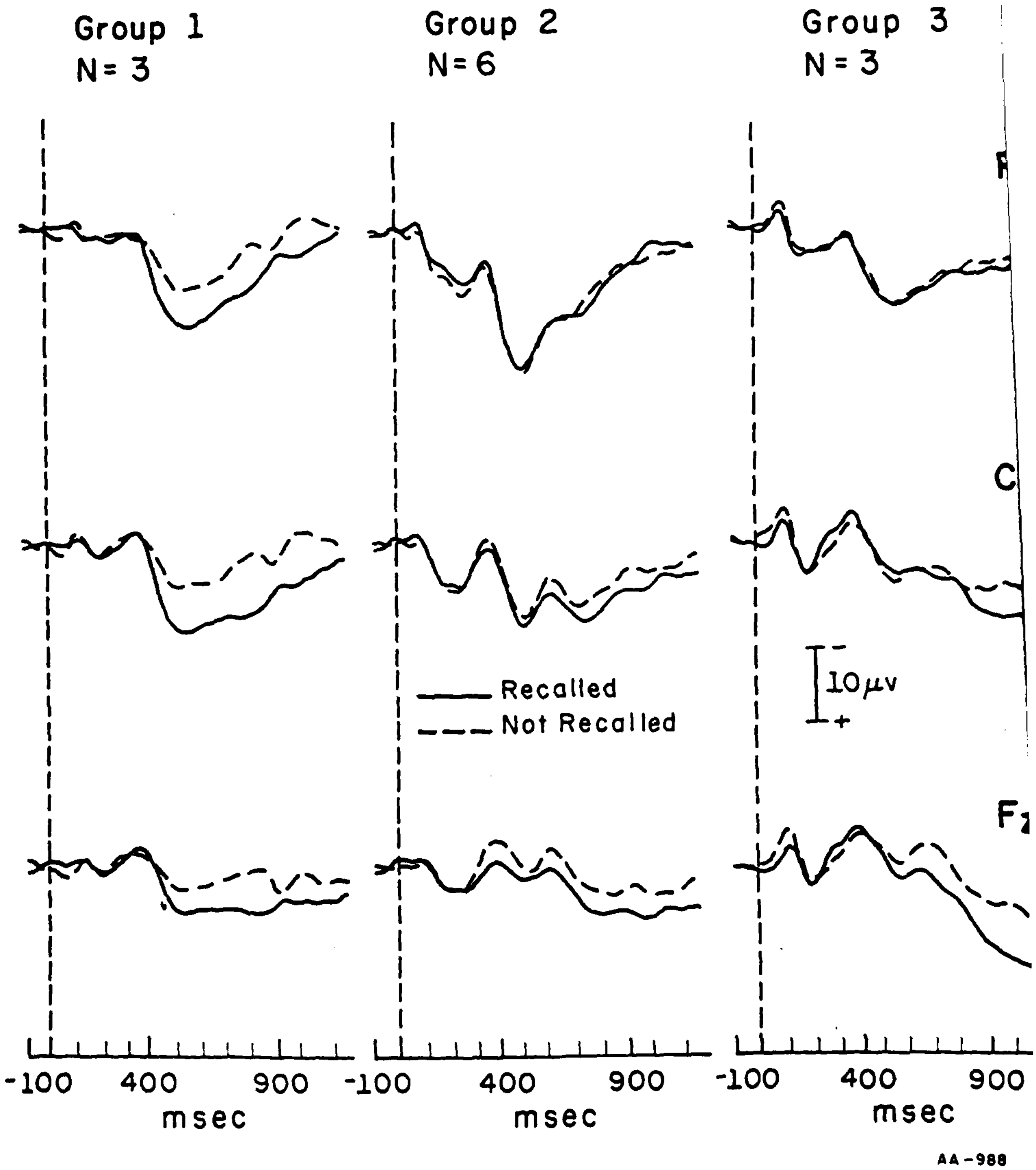

Group 1
N= 3
Group 2
N=6
Group 3
N= 3
C
F
Recalled
Not Recalled
10μv
-100
400
900
msec
AA-988

Figure 7

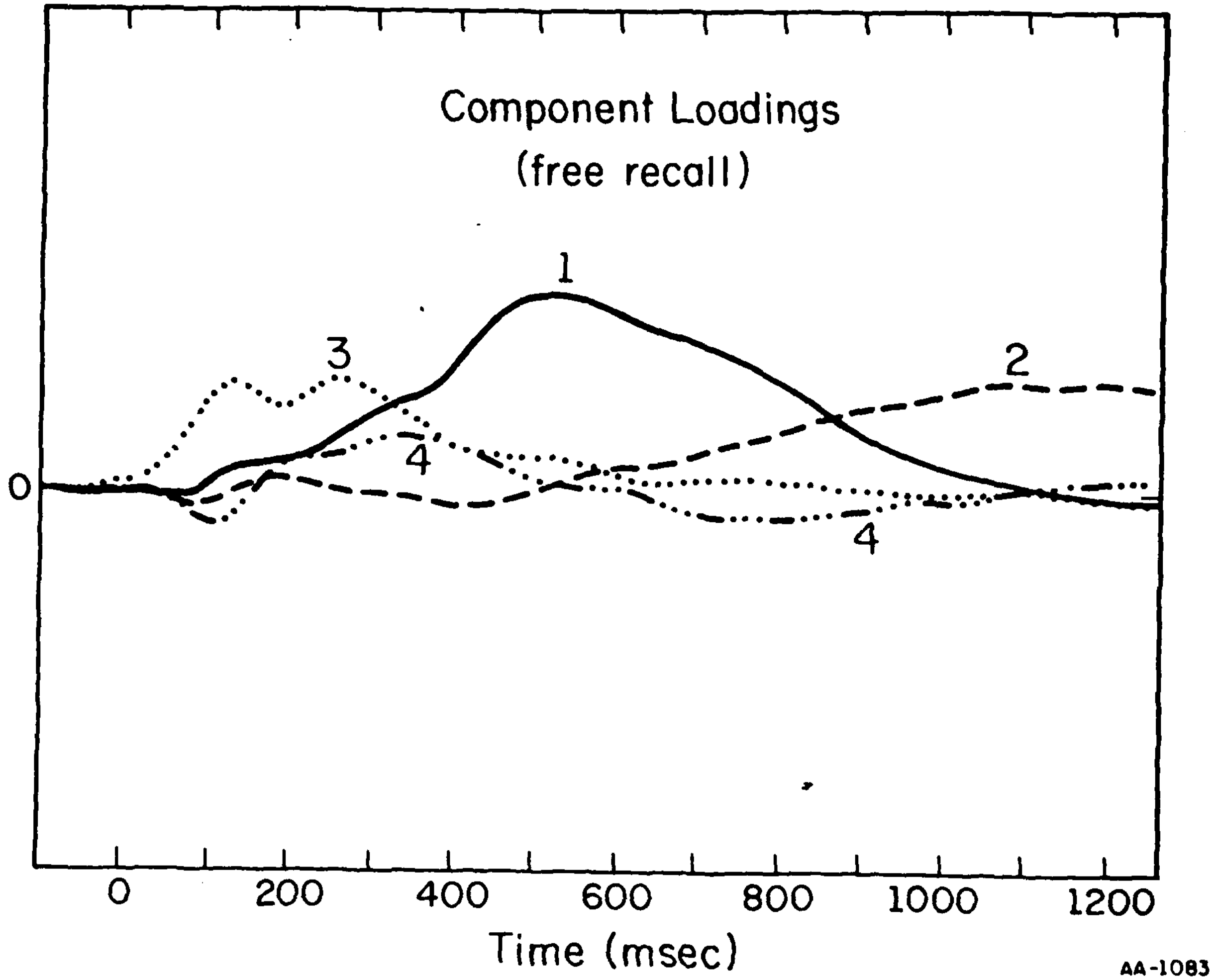
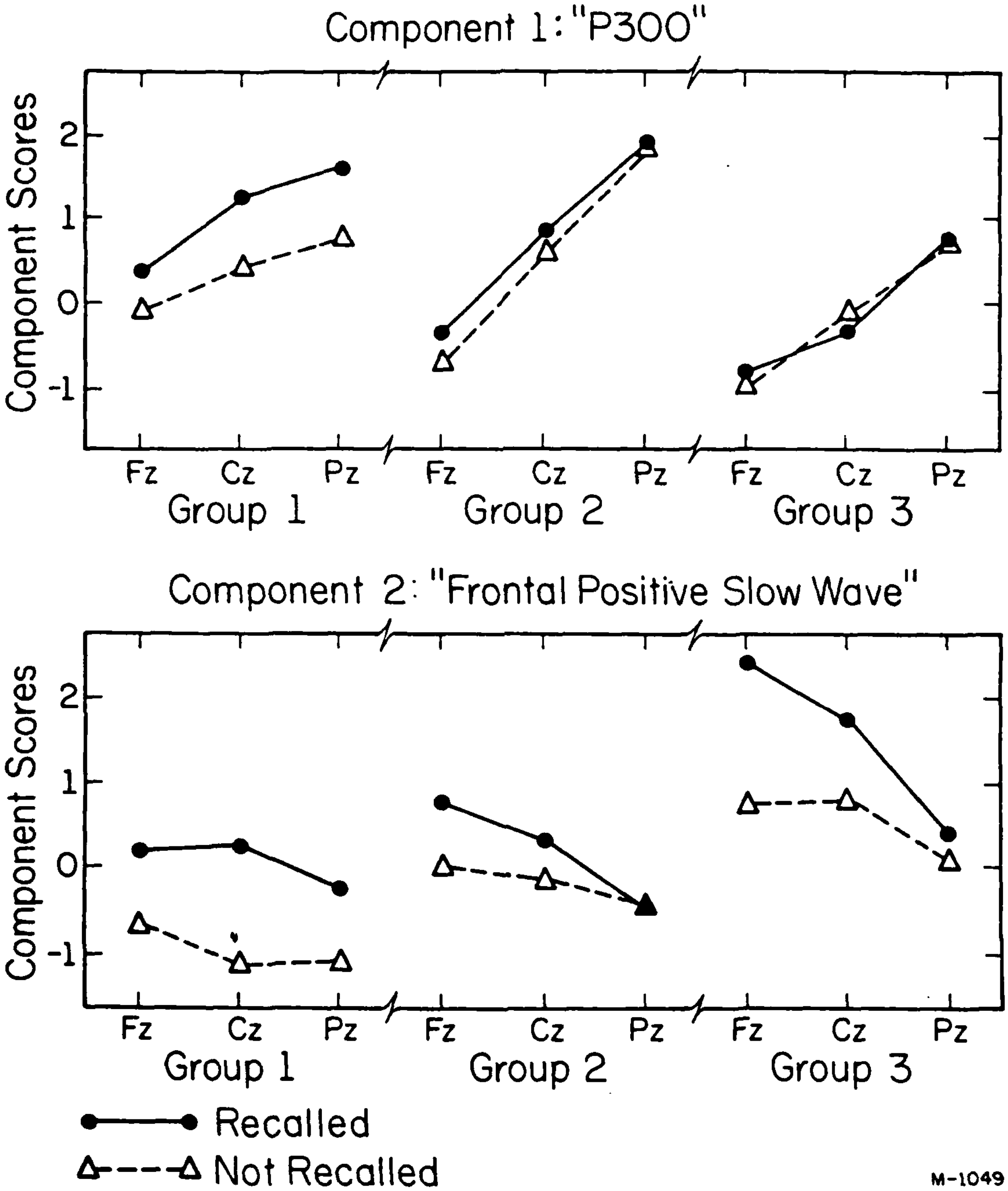
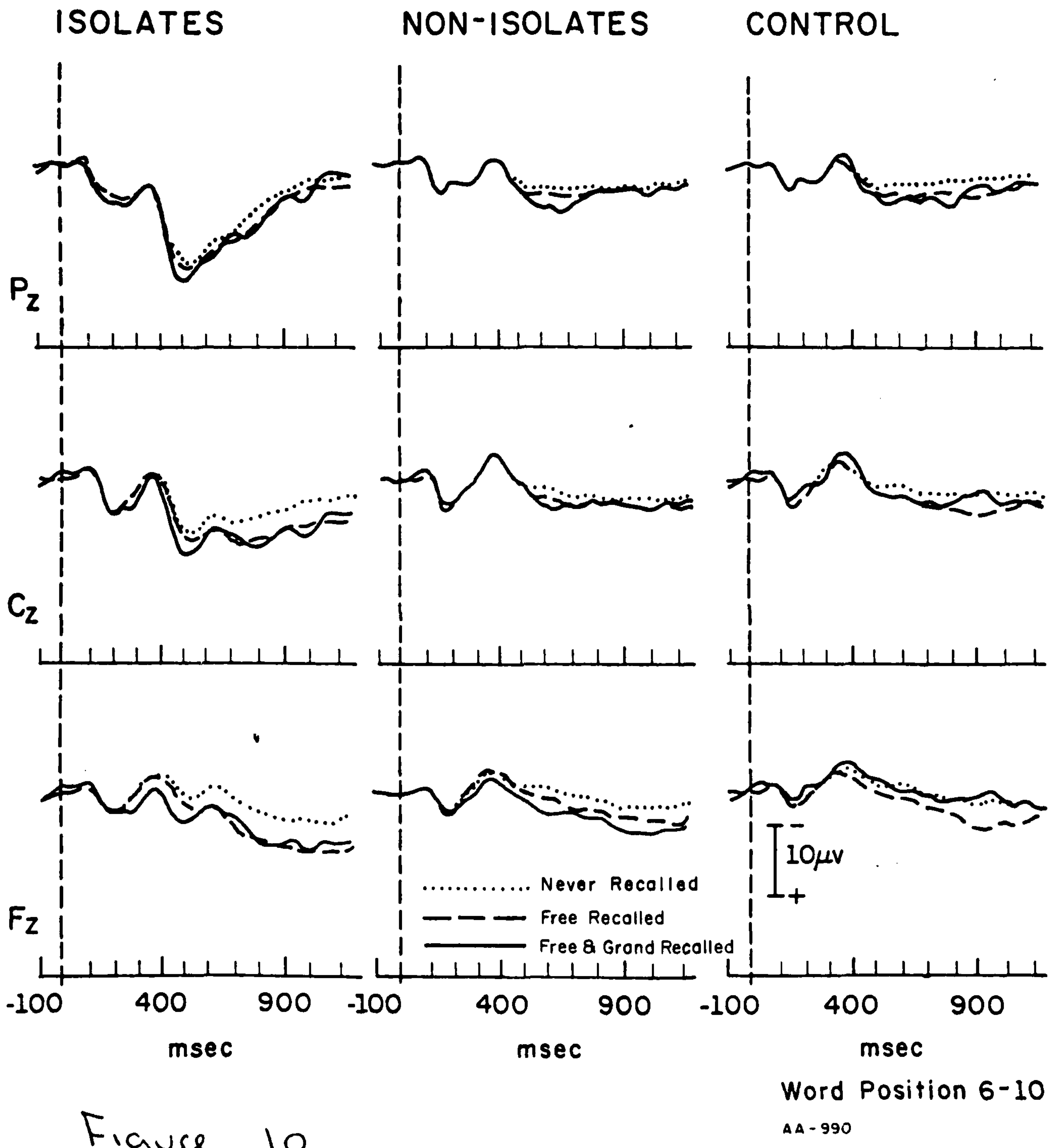

Component Loadings
(free recall)
1
2
3
4
4
0
0
200
400
600
800
1000
1200
Time (msec)
AA-1083

Figure 8



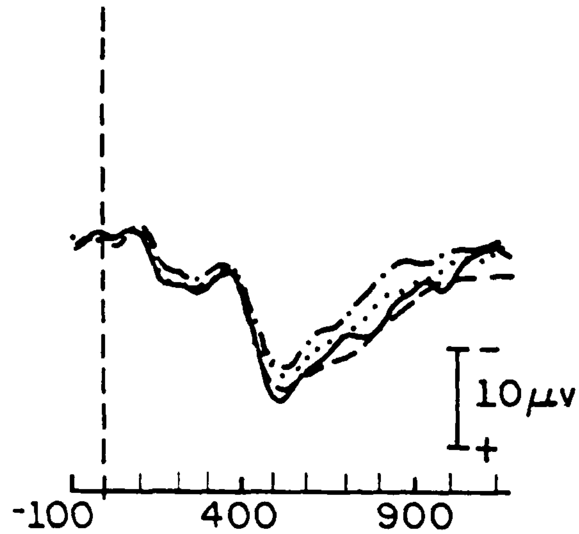
Component 1: "P300"
Component Scores
2
1
0
-1
Fz Cz Pz Fz Cz Pz Fz Cz Pz
Group 1 Group 2 Group 3
Component 2: "Frontal Positive Slow Wave"
Component Scores
2
1
0
-1
Fz Cz Pz Fz Cz Pz Fz Cz Pz
Group 1 Group 2 Group 3
Recalled
Not Recalled
M-1049

Figure 9



ISOLATES
NON-ISOLATES
CONTROL
Pz
Cz
Fz
10μv
Never Recalled
Free Recalled
Free & Grand Recalled
-100
400
900
msec
Word Position 6-10
AA-990

Figure 10



Recognized, and Recalled in Both Free and Grand Recall
Recognized and Recalled in Free Recall Only
Recognized Only
Neither Recognized Nor Recalled

AA-989

Figure 11

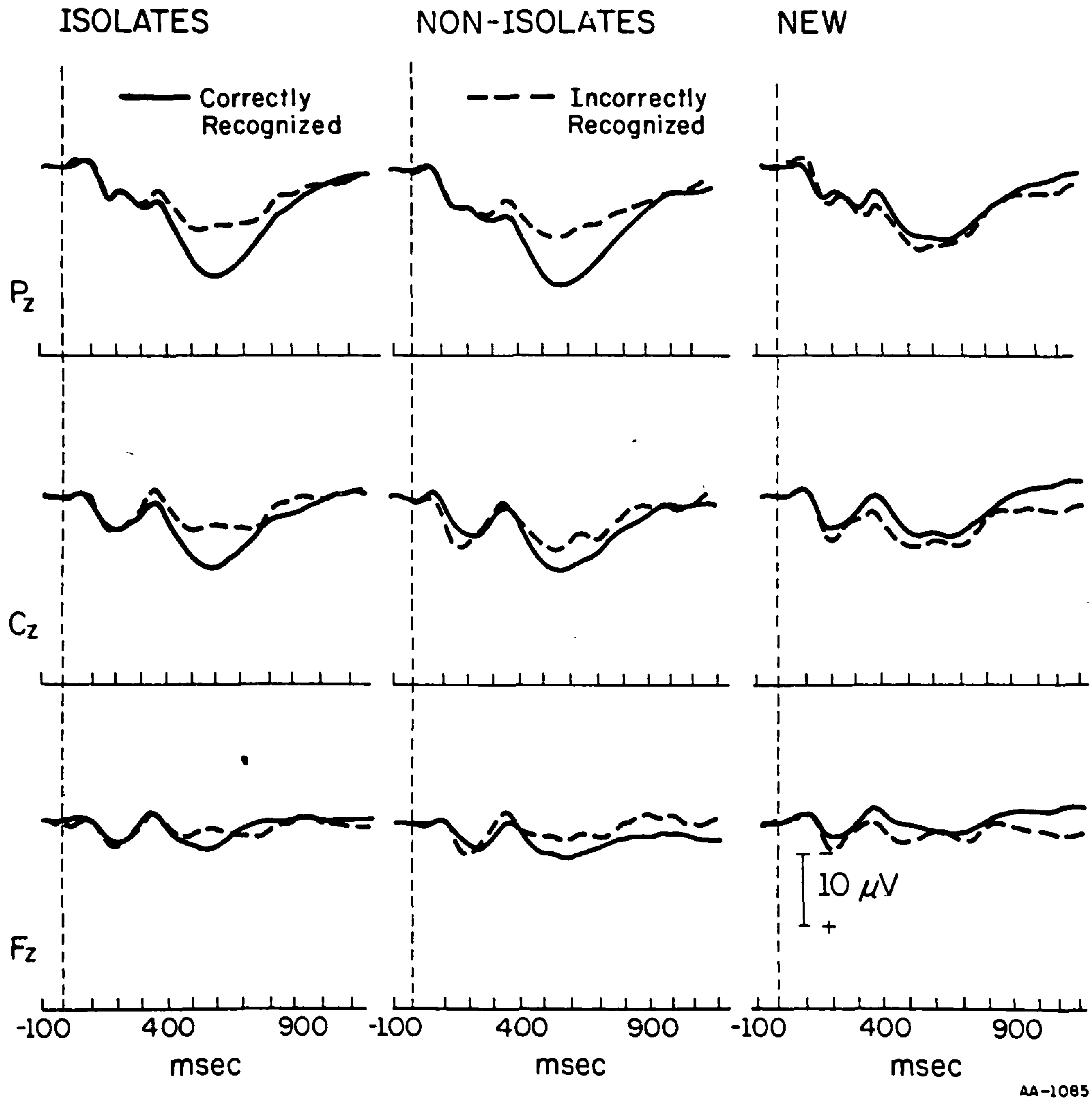

ISOLATES
NON-ISOLATES
NEW
Correctly Recognized
Incorrectly Recognized
Pz
Cz
Fz
10 μV
+
-100
400
900
msec
AA-1085

Figure 12

P300 and Recall in an Incidental Memory Paradigm

M. Fabiani, D. Karis, M.G.H. Coles, and E. Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

In previous work using a free recall paradigm (Fabiani, Karis, & Donchin, 1982) we found striking individual differences in recall, and also in the relationship between the ERPs elicited by words and the later recall of those words. For subjects who used rote mnemonic strategies, words that were later recalled elicited larger amplitude P300s on their initial presentation than words that were not recalled. Subjects who used elaborative strategies did not show this relationship. We argued that for these subjects the relationship between P300 and recall was overshadowed by the mnemonically powerful associative processing that continued long after the time period reflected by P300. In the present study we assessed the relationship between P300 amplitude and recall in an incidental memory paradigm in which recall was not expected at the time the material was presented.

We created such a paradigm by embedding a free recall test in a series of five "oddballs". In the fourth oddball each of 12 male subjects was presented with a random sequence of 105 male and female names. Each name was presented only once. Subjects were instructed to count either the rare (n=21) or the frequent (n=84) names and report a running total at the end. Immediately afterwards they were given five minutes to write down as many names (male and female) as they could remember. This recall was unexpected and all subjects expressed surprise.

One name was presented every 2 seconds and the ERPs elicited by each name were recorded from Fz, Cz, and Pz (referred to linked mastoids) using an eight second time constant and an upper half amplitude cutoff of 35 Hz. EEG and EOG were digitized at 100 samples/sec for 150 points beginning 100 msec prior to the presentation of a name. Eye movement artifacts were corrected off-line using a procedure described in Gratton, Coles, and Donchin (1983). ERP averages were computed for each subject by sorting according to recall, and the difference between baseline and the most positive peak between 250 and 1000 msec (at Pz) was chosen as the P300 amplitude. In 10 of the 12 subjects larger amplitude P300s were elicited by words that were recalled than by words that were not recalled. This relationship was confirmed ($p<.01$) by an analysis of variance on the amplitude values. It supports our theory that P300 reflects processes invoked when events occur and create a need to revise the current representations in working memory ("context updating"; Donchin, 1981).

P300 and Recall in an Incidental Memory Paradigm

M. Fabiani, D. Karis, M. G.H. Coles, and E. Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

In previous work using a free recall paradigm (Fabiani, Karis, & Donchin, 1982) we found striking individual differences in recall, and also in the relationship between the ERPs elicited by words and the later recall of those words. For subjects who used rote mnemonic strategies, words that were later recalled elicited larger amplitude P300s on their initial presentation than words that were not recalled. Subjects who used elaborative strategies did not show this relationship. We argued that for these subjects the relationship between P300 and recall was overshadowed by the mnemonically powerful associative processing that continued long after the time period reflected by P300. In the present study we assessed the relationship between P300 amplitude and recall in an incidental memory paradigm in which recall was not expected at the time the material was presented.

We created such a paradigm by embedding a free recall test in a series of five "oddballs". In the fourth oddball each of 12 male subjects was presented with a random sequence of 105 male and female names. Each name was presented only once. Subjects were instructed to count either the rare (n=21) or the frequent (n=84) names and report a running total at the end. Immediately afterwards they were given five minutes to write down as many names (male and female) as they could remember. This recall was unexpected and all subjects expressed surprise.

One name was presented every 2 seconds and the ERPs elicited by each name were recorded from Fz, Cz, and Pz for 1.5 seconds. ERP averages were computed for each subject by sorting according to recall, and the difference between baseline and the most positive peak between 250 and 1000 msec (at Pz) was chosen as the P300 amplitude. In 10 of the 12 subjects larger amplitude P300s were elicited by words that were recalled than by words that were not recalled. This relationship was confirmed ($p<.01$) by an analysis of variance on the amplitude values. It supports our theory that P300 reflects processes invoked when events occur and create a need to revise the current representations in working memory ("context updating"; Donchin, 1981).

APPENDIX A2b

Poster presented at conference of Society for Psychophysiological Research, September 1983, Asilomar Conference Center, Pacific Grove, California

P300 and Recall in an Incidental Memory Paradigm

M. Fabiani, D. Karis, M.G.H. Coles, and E. Donchin

Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

INTRODUCTION

In previous work using a free recall paradigm (Fabiani, Karis, & Donchin, 1982) we found striking individual differences in recall, and also in the relationship between the ERPs elicited by words and the later recall of those words. For subjects who used rote mnemonic strategies, words that were later recalled elicited larger amplitude P300s on their initial presentation than words that were not recalled. Subjects who used elaborative strategies did not show this relationship. We argued that for these subjects the relationship between P300 and recall was overshadowed by the mnemonically powerful associative processing that continued long after those processes reflected by P300 had terminated.

Therefore we hypothesize that if individual differences due to mnemonic strategies are suppressed, the relationship between P300 amplitude and recall should hold for most of the subjects. To test this hypothesis we used an incidental memory paradigm in which recall was not expected at the time the material was presented. In this situation subjects are unlikely to engage in elaborate rehearsal, and individual differences in encoding and rehearsal should be minimized.

METHOD

We created an incidental memory paradigm by embedding a free recall test in a series of five "oddballs". The experimental design is depicted in Figure 1. In the fourth oddball each of 35 male subjects was presented with a random sequence of 105 male and female names. Each name was presented only once. Subjects were instructed to count either the rare (n=21) or the frequent (n=84) names and report a running total at the end. Immediately afterwards they were given five minutes to write down as many names (male and female) as they could remember. This recall was unexpected and all subjects expressed surprise.

One name was presented every 2 seconds and the ERPs elicited by each name were recorded from Fz, Cz, and Pz (referred to linked mastoids) using an 8 second time constant and an upper half amplitude cutoff of 35 Hz. EEG and EOG were digitized at 100 samples/sec for 150 points beginning 100 msec prior to the presentation of a name. Eye movement artifacts were corrected off-line using a procedure described by Gratton, Coles, and Donchin (1983).

DATA ANALYSIS

ERP averages were computed for each subject and each oddball (count rare and count frequent) according to type of stimulus (rare or frequent). For the first name oddball (which was followed by the incidental free recall test), ERP averages were also computed for each subject by sorting the trials according to recall (recalled or not recalled in the subsequent test). Given the latency variability observed among subjects, average

waveforms sorted on the basis of recall were latency adjusted for eacn subject and condition. A cross-correlation procedure was used to estimate P300 latency (window from 350 to 800 msec). The template adopted was a 2 Hz cosinusoidal wave (1 cycle). P300 amplitude was assessed by means of a peak-to-peak procedure, the slope at the maximal cross-correlation function (b=r*Sy/Sx; where Sx is the variance of the template and is constant). This measure is the least-square estimate of the waveform to template ratio over the entire window. This procedure was chosen in order to minimize noise due to the small number of trials in each average and because of its insensitivity to errors in baseline definition.

RESULTS

Name oddballs

Subjects were generally very accurate in the count task. Only four subjects were more than one off the correct count in one of the two name oddballs.

Average ERPs for rare and frequent names in each oddball (count rare and count frequent) are presented in Figure 2 for both groups of subjects. As expected, the rare names elicited larger P300s than the frequent names. However this probability effect was tempered by a large "target effect": the counted names elicited larger P300s than the uncounted names. Target rare names showed the largest P300 and non-target frequent names the smallest.

It is also noteworthy that ERP's recorded under comparable instructions (count rare - count frequent) for the two groups are very similar, even

though the groups were given the instructions in reverse order.

Memory Results

Subjects recalled more counted (target) names than non-counted (non-target) names, as shown in Figure 3.

ERPs for rare and frequent names were divided into two classes - those later recalled and those not recalled. The latency adjusted and unadjusted grand averages are shown in Figures 4-7.

An ANOVA was performed on the amplitude estimates derived using the procedure described above. The following results were significant with $p<0.05$:

1. Names recalled in the subsequent test showed a larger amplitude P300 than names not recalled ($F=14.13$; $df=1,33$).

2. The difference in P300 amplitude between names recalled and not recalled was larger for the rare names than for the frequent ($F=5.82$; $df=1,33$).

Performance and the Memory Effect

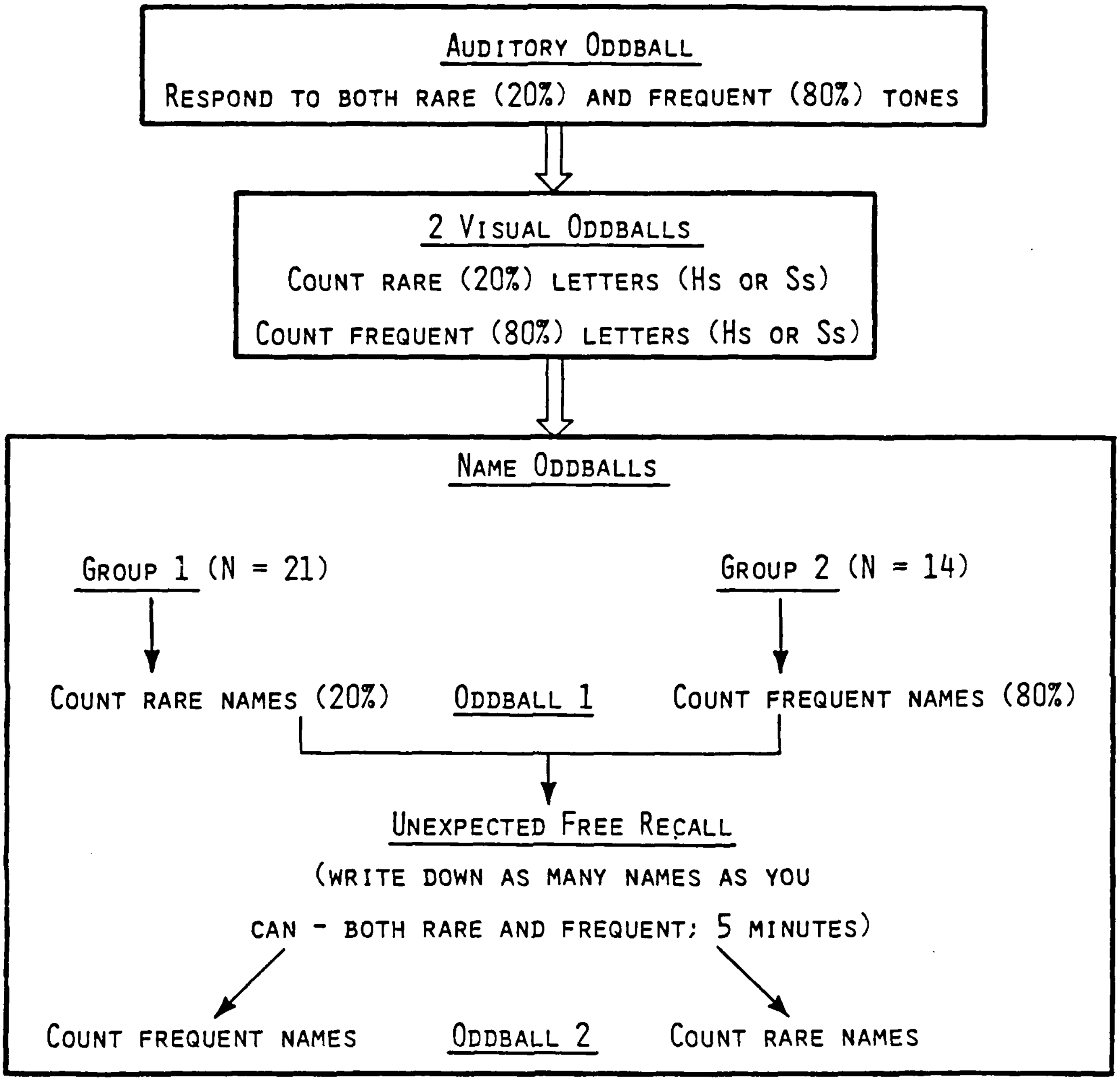
For the counted names an interesting relationship between performance and the memory effect emerged: the memory effect (the difference in P300 amplitude between recalled and not-recalled target names) was negatively correlated with recall performance. That is, the more counted names a

subject recalled, the smaller the difference in P300 amplitude between recalled and not recalled counted names ($r=-.49$, for subjects who counted rare; $r=-.64$, for subjects who counted frequent; $p<.05$). These negative correlations between the memory effect and performance may result from strategy differences during retrieval. Some subjects, for example, reported going through the alphabet or thinking of common names, and then trying to decide if those names had been presented. In these cases, recall may depend more on the effectiveness of the retrieval strategies than on the extent of "context updating" indexed by P300 amplitude.

CONCLUSIONS

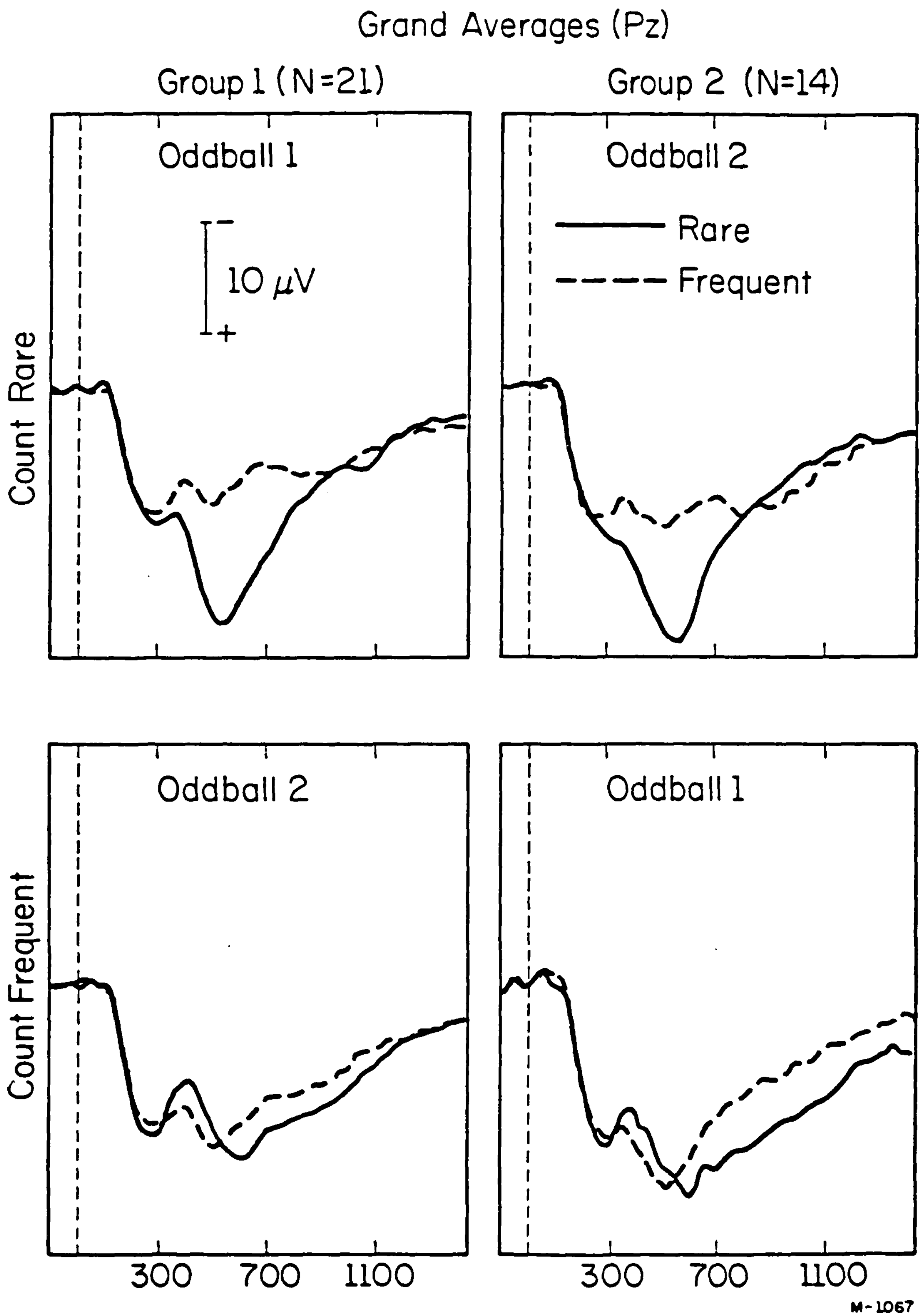
The results support our theory that P300 reflects processes invoked when events occur and create a need to revise the current representations in working memory ("context updating"; Donchin, 1981). In fact, when subjects are not likely to use mnemonic strategies to memorize materials, a strong relationship between P300 amplitude and subsequent recall emerges. We think we have been able to minimize the individual differences that can emerge during item encoding and rehearsal. However, differences will always remain during retrieval, and they may also influence the relationship between P300 and recall.

EXPERIMENTAL DESIGN

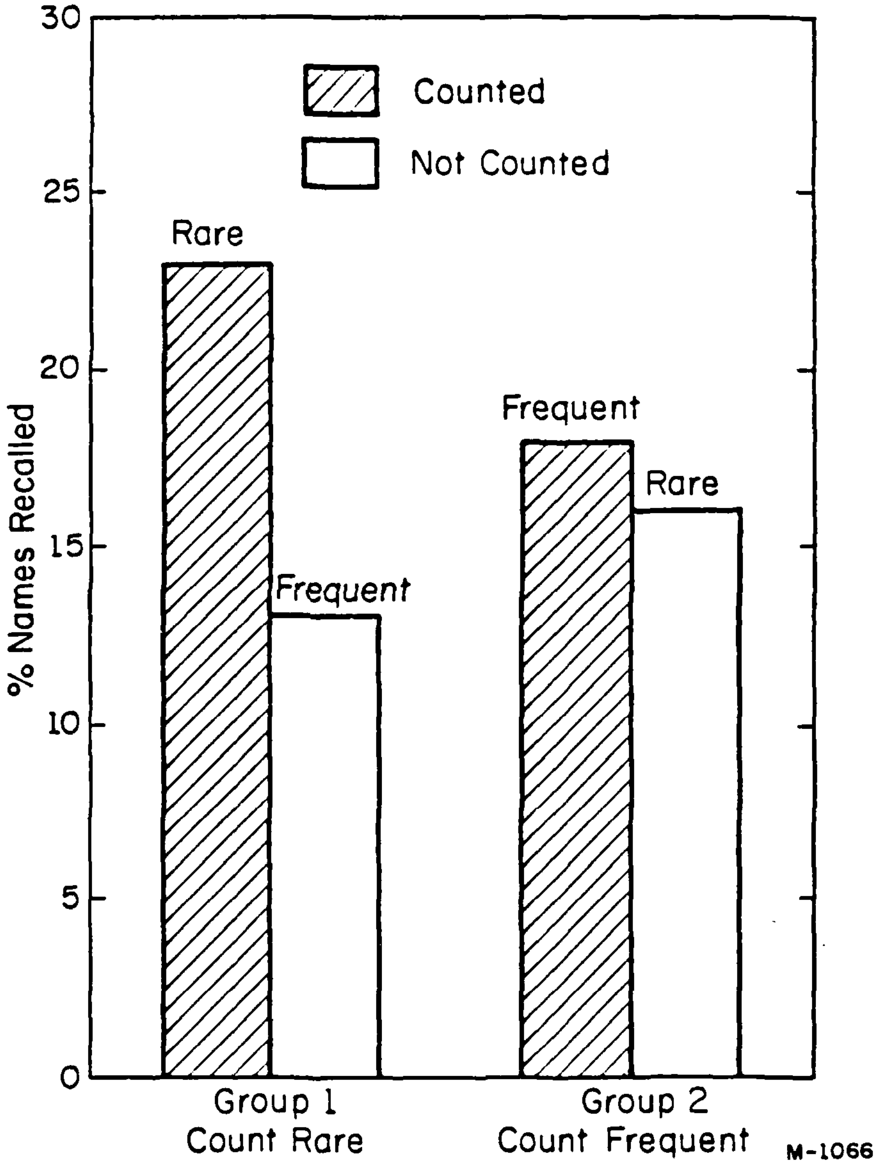


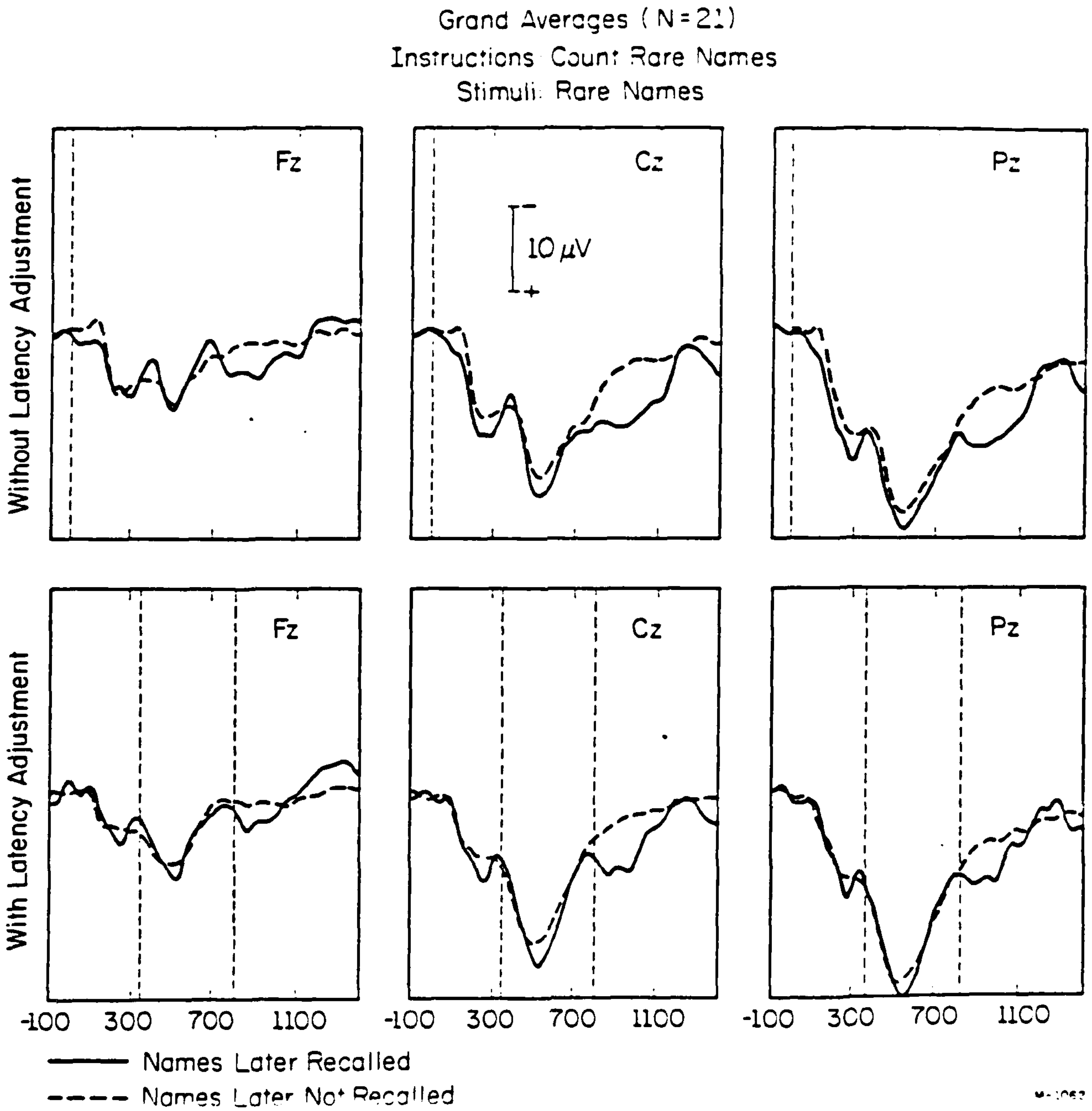
105 NAMES IN EACH ODDBALL
ISI = 2 SECONDS
STIMULUS DURATION = 200 MSEC

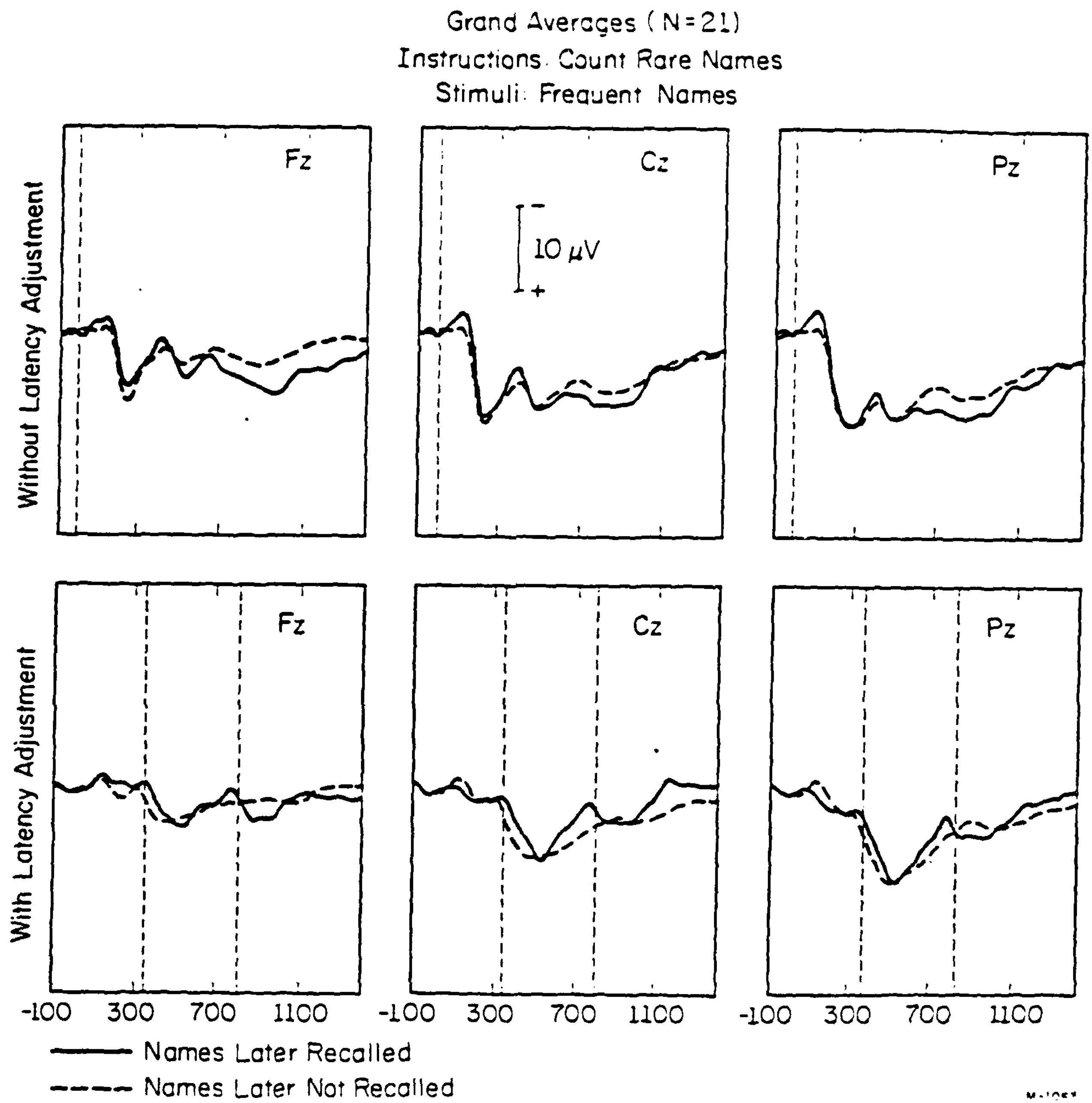
IN EACH GROUP MALE NAMES WERE RARE FOR HALF THE SUBJECTS.

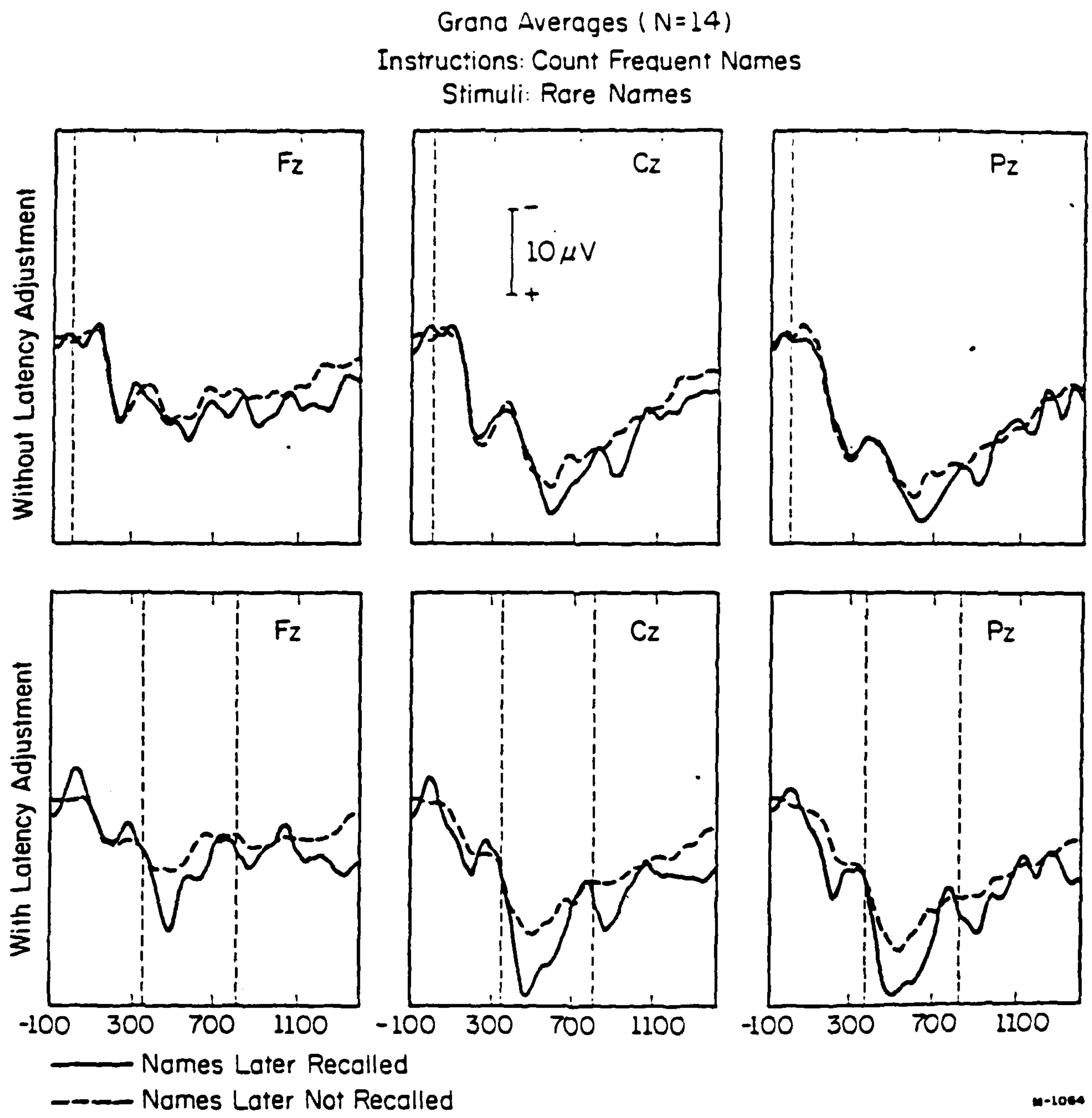


Grand Averages (Pz)
Group 1 (N=21)
Group 2 (N=14)
Oddball 1
Oddball 2
10 μV
Rare
Frequent
Count Rare
Count Frequent
Oddball 2
Oddball 1
300
700
1100
300
700
1100
M-1067

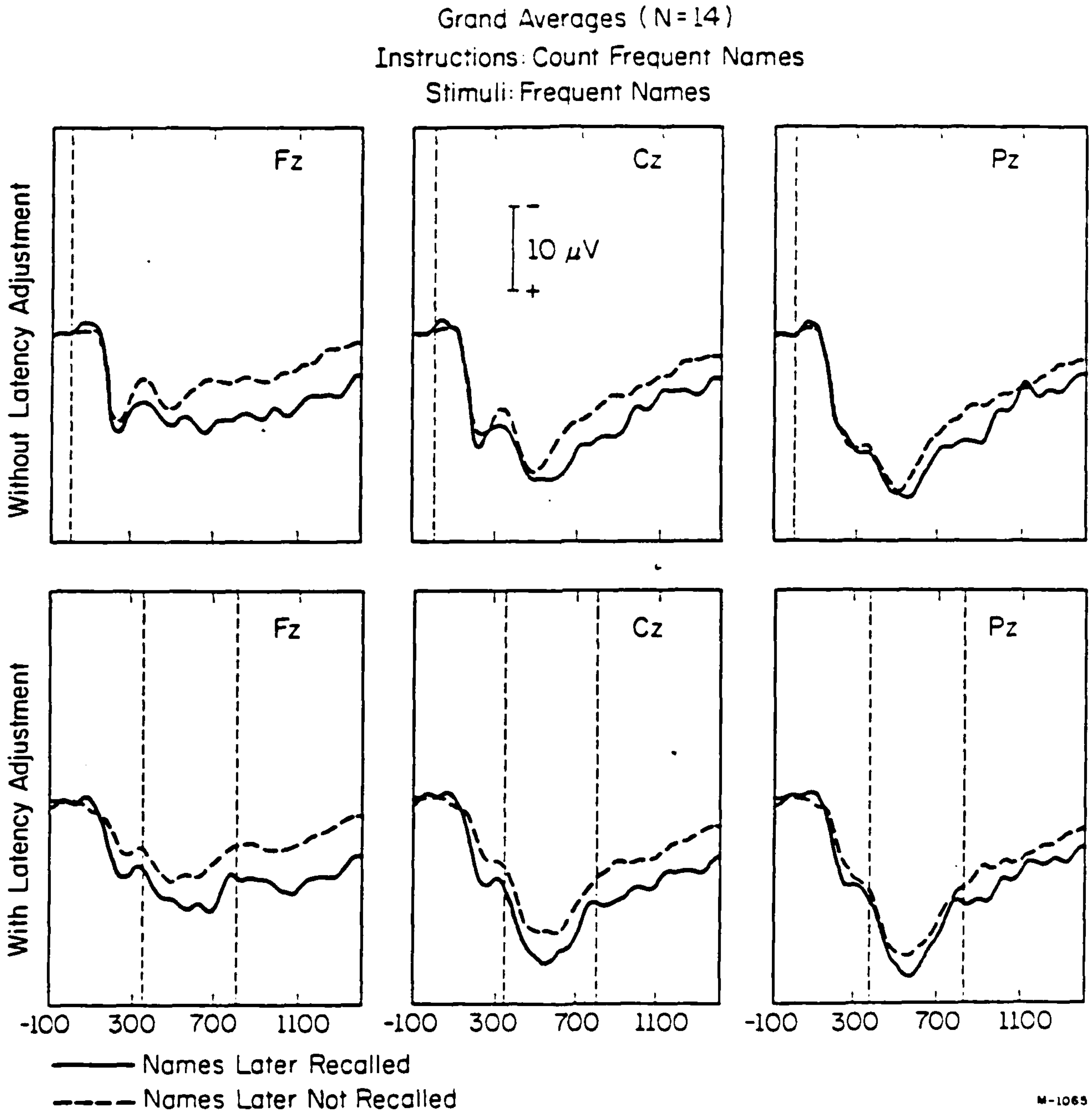

30
25
20
15
10
5
0
% Names Recalled
Counted
Not Counted
Rare
Frequent
Frequent
Rare
Group 1
Count Rare
Group 2
Count Frequent
M-1066


Grand Averages (N=21)
Instructions: Count Rare Names
Stimuli: Rare Names
Without Latency Adjustment
With Latency Adjustment
Fz
Cz
Pz
10 μV
-
+
-100
300
700
1100
Names Later Recalled
Names Later Not Recalled


Grand Averages (N=21)
Instructions: Count Rare Names
Stimuli: Frequent Names
Without Latency Adjustment
With Latency Adjustment
Fz
Cz
Pz
10 μV
-100
300
700
1100
Names Later Recalled
Names Later Not Recalled



Grand Averages (N=14)
Instructions: Count Frequent Names
Stimuli: Rare Names
Without Latency Adjustment
With Latency Adjustment
Fz
Cz
Pz
10 μV
-100
300
700
1100
Names Later Recalled
Names Later Not Recalled
M-1064


Grand Averages (N=14)
Instructions: Count Frequent Names
Stimuli: Frequent Names
Without Latency Adjustment
With Latency Adjustment
Fz
Cz
Pz
10 µV
-100
300
700
1100
Names Later Recalled
Names Later Not Recalled
M-1065

Paper to be presented at the [illegible]
Society Meeting, June 29th - July 1st, 1983, Bristol, England

P300 and Response Accuracy: An Analysis Using Response Bias and Error Titration

Michael G. H. Coles, Gabriele Gratton, David Dupree,
Theodore R. Bashore, Charles W. Eriksen, and Emanuel Donchin

Cognitive Psychophysiology Laboratory - Department of Psychology
University of Illinois, Champaign, Illinois 61820, USA

Kutas, McCarthy and Donchin (1977) found that reaction time (RT) was shorter and P300 latency was longer when subjects erred in a choice RT task. Two experiments were designed to evaluate the P300/Error relationship in more detail by (a) manipulating response bias and (b) "titrating" the degree of error. In the first, 7 subjects performed a choice RT task (male versus female names) under speed instructions with unequal probability of the two response classes (.2 and .8). EEG from Fz, Cz, and Pz, RT, and response accuracy were derived for each trial. The RT distribution for rare error trials was the same as that for correct frequent trials. However, P300 latency was longer on rare error trials than on both correct frequent and correct rare trials. In the second experiment, 12 subjects performed a choice RT task (letter "H" or "S", probability of .5). Measures were as above, plus EMG and force activity for the two responding hands. By evaluating EMG and force measures on both correct and incorrect sides on each trial, it was possible to define a "degree of error" dimension. As the degree of error increased the latency of both P300 and correct activity increased, while that for incorrect activity decreased. These results indicate that when incorrect activity is present both the correct activity and P300 are delayed (or inhibited). Whether the increase in P300 latency indicates a dependence of P300 on the recognition of an error or merely that stimulus evaluation processes are longer on error trials remains to be determined.

Kutas, M., McCarthy, G., and Donchin, E., Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. Science, 1977, 197, 792-795

Figure 1.

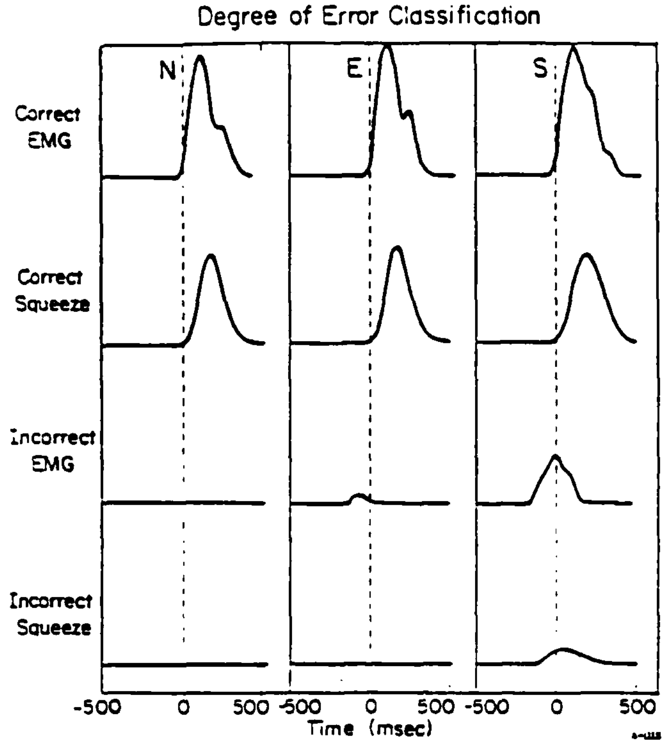
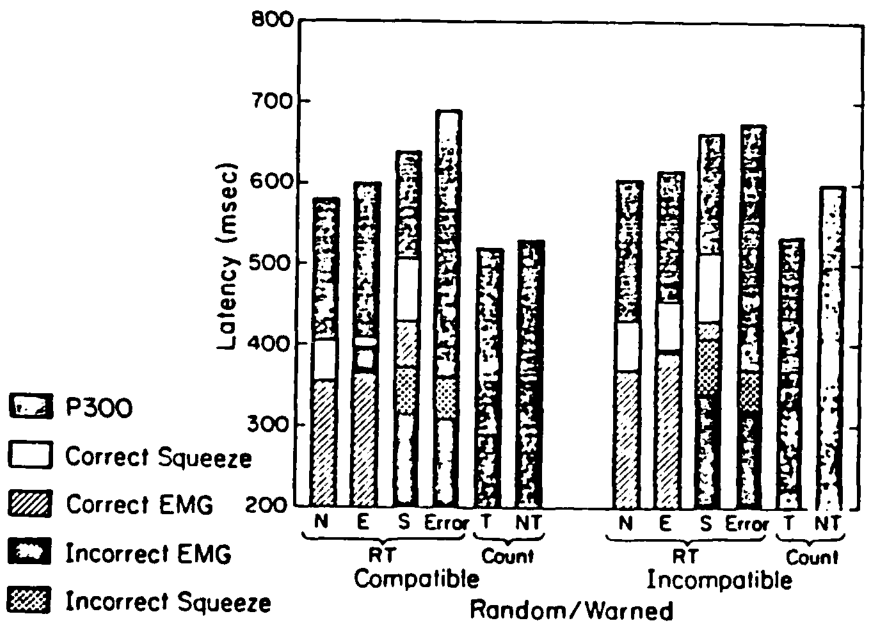


Figure 1. Analog representation of the degree of error classification. The N category has only EMG and squeeze activity on the correct side. E and S categories also have EMG, and EMG and squeeze activity, on the incorrect side. The Error category (not shown) has EMG and squeeze activity on the incorrect side only.

Figure 2. Latencies of P300 and correct and incorrect EMG and squeeze activity as a function of error classification and compatibility for the random (variable)/warned condition. Latencies of P300 for the count task are also shown.

Figure 2.



An ERP/EMG/RT Approach to the Continuous Flow Model of Cognitive Processes

Gabriele Gratton, Michael G. H. Coles, Theodore R. Bashore,
Charles W. Eriksen, and Emanuel Donchin

Cognitive Psychophysiology Laboratory - Department of Psychology
University of Illinois, Champaign, Illinois 61820, USA

The continuous flow model of cognitive processing applied to visual search proposes that "information accumulates gradually in the visual system, with concurrent priming of responses" (Eriksen & Schultz, 1979). To test this model, the present experiment used measures of the P300 component of the event-related potential to assess the duration of the stimulus evaluation process, and measures of the electromyogram and response force to titrate response processes.

Twelve male students received 8 blocks of 80 trials each of a discrimination task. On each trial, they were required to respond to target letters "H" or "S" by squeezing a zero displacement dynamometer with the left or right hand. The target letter, which was presented at the visual fixation point, was embedded in a set of either compatible (e.g. HHHHH) or incompatible (e.g. SSHSS) letters. The level of compatibility was either variable or fixed within a trial block (blocking manipulation), and, for half the blocks, a warning tone preceded target letter presentation by 1 sec. Measures of EEG (Fz, Cz, and Pz), EOG, EMG from each forearm, and squeeze force for each hand, were obtained in analog form and digitized on-line at 100 Hz. For each trial, latency measures were derived off-line for the onset of EMG and squeeze activity (RT), and for P300. Then, each trial was classified into one of four categories on the basis of the EMG and squeeze activity for correct and incorrect sides (see Figure 1). This classification system yielded a "degree of error" dimension.

Trials for which there was some evidence of incorrect activity were more common under incompatible than compatible conditions. Latencies of all measures were shorter under conditions of compatibility and blocking. EMG and squeeze activity on both sides, but not P300, occurred earlier in the warning condition. Latency of correct activity and P300 increased, while that of incorrect activity decreased, with degree of error (see Figure 2).

Data were also obtained for a control condition, in which subjects merely counted one of the target letters. As in the RT task, P300 latency was influenced by compatibility, although it was consistently shorter in the count task.

These data suggest that the degree of error and latency of correct activity are a function of (a) an activation process that is independent of the nature of the stimulus, (b) the rate of accumulation of evidence for a particular target stimulus provided by an evaluation process, and (c) a response interference mechanism. This interpretation is in accordance with a continuous flow model of cognitive processes.

Eriksen, C. W., & Schultz, D. W. Information processing in visual search: A continuous flow conception and experimental results. Perception and Psychophysics , 1979, 25, 249-263.

P300, Practice and Consistency in Visual Search

Arthur Kramer, Arthur Fisk and Walter Schneider

Cognitive Psychophysiology Laboratory
and Human Performance Laboratory

University of Illinois

The two process theory proposed by Schneider and Shiffrin (1977) provides an interpretation of the qualitative differences in processing that occur with extended practice. Automatic processing typically develops when subjects deal with consistent stimulus-response mapping over many trials (CM condition). The automatic processing mode is characterized as fast, inflexible, capable of being performed in parallel with other tasks and insensitive to the number of items to be maintained in memory or the number of items in a display. The controlled proccessing mode is employed when subjects are unable to consistently map stimuli to responses (varied mapping (VM) condition). Controlled processing is slow, serial and resource sensitive.

The present study focused on the effects of, and the interactions between, practice and task structure on human performance. The development of the automatic mode (CM condition) was assessed by means of measures of reaction time (RT) and event-related brain potentials (ERP). It was hypothesized that RT and P300 latency would increase linearly with the number of items to compare (memory set) in the VM condition but not in the practiced CM condition. Furthermore, it was expected that the commonly observed relation between subjective probability and P300 amplitude, larger P300s elicited by infrequent events, would be attenuated in the CM condition after extensive practice. The P300 probability effect has been suggested to be the result of memory updating and therefore should be unnecessary during the automatic mode. The paradigm employed to investigate these issues was similar to the modified Sternberg task (1969) used by Schneider and Shiffrin (1977).

Each trial began with a 10 sec presentation of the memory set (1 or 4 items). In the 30 frames which followed the presentation of each memory set, the subjects task was to press a button if a memory set item (target) was present (go-nogo task). Three variables were orthogonally manipulated in a factorial design. These variables included the number of memory set items (1 or 4), the type of training (CM or VM) and the probability of occurrence of a memory set item (.20 or .80). In the Cm condition targets were always selected from one category (numbers 1 to 9) while distractors were choosen from another category (letters A to I). In the VM condition both the targets and distractors were choosen from the same category (letters A to I). Targets and distractors exchanged roles over trials in the VM conditions. Twelve sessions of 1680 trials were run with each of 5 subjects. RT's and accuracy measures were obtained in all of the sessions. ERP's were recorded in the first and twelfth sessions. Three channels (Fz, Cz and Pz) and vertical EOG were digitized at 100 Hz over epochs extending 100 msec before and 1700 msec subsequent to the presentation of a target or distractor.

RT results were consistent with other research employing similar paradigms. Set size had a significant effect on RT in both CM and VM conditions in session 1 and the VM condition in session 12. However, the set size effect was not significant in the CM condition after twelve sessions. Error rate was not influenced by experimental manipulations (less than 2%). P300 latency mirrored RT suggesting that the development of automatic processing substantially reduced stimulus evaluation time. The

effect of probability on P300 amplitude obtained for both CM and VM conditions in session 1 was not significant for the CM condition in session 12. This finding suggests an attenuation of memory updating with CM practice. The amplitude of the nogo P300s was reduced in comparison to the go P300s. This effect may be attributed to a larger overlapping negative component in the nogo condition.

P300, Consistency and Visual Search

Arthur Kramer, Arthur Fisk and Walter Schneider

Department of Psychology
University of Illinois
Champaign, Illinois

Presented at the Twenty-Third Annual Meeting of
The Society for Psychophysiological Research (1983)

INTRODUCTION

THE TWO PROCESS THEORY PROPOSED BY SCHNEIDER AND SHIFFRIN (1977) PROVIDES AN INTERPRETATION OF THE QUALITATIVE DIFFERENCES IN PROCESSING THAT OCCUR WITH EXTENDED PRACTICE. AUTOMATIC PROCESSING TYPICALLY DEVELOPS WHEN SUBJECTS DEAL WITH CONSISTENT STIMULUS-RESPONSE MAPPING OVER MANY TRIALS (CM CONDITION). THE AUTOMATIC PROCESSING MODE IS CHARACTERIZED AS FAST, INFLEXIBLE, IS CAPABLE OF OCCURING IN PARALLEL WITH OTHER TASKS, AND IS INSENSITIVE TO THE NUMBER OF ITEMS TO BE MAINTAINED IN MEMORY OR THE NUMBER OF ITEMS IN A DISPLAY. THE CONTROLLED PROCESSING MODE IS EMPLOYED WHEN SUBJECTS ARE UNABLE TO CONSISTENTLY MAP STIMULI TO RESPONSES (VM CONDITION). CONTROLLED PROCESSING IS SLOW, SERIAL AND RESOURCE SENSITIVE.

THE PRESENT STUDY FOCUSED ON THE EFFECTS OF, AND THE INTERACTIONS BETWEEN, PRACTICE AND TASK STRUCTURE ON HUMAN PERFORMANCE. THE DEVELOPMENT OF THE AUTOMATIC MODE (CM CONDITION) WAS ASSESSED BY MEANS OF MEASURES OF REACTION TIME (RT) AND EVENT-RELATED BRAIN POTENTIALS (ERP). IT WAS HYPOTHESIZED THAT RT AND P300 LATENCY WOULD INCREASE LINEARLY WITH THE NUMBER OF ITEMS TO COMPARE (MEMORY SET) IN THE VM CONDITION BUT NOT IN THE PRACTICED CM CONDITION. FURTHERMORE, IT WAS EXPECTED THAT THE COMMONLY OBSERVED RELATION BETWEEN SUBJECTIVE PROBABILITY AND P300 AMPLITUDE (LARGER P300S ELICITED BY INFREQUENT EVENTS) WOULD BE ATTENUATED IN THE CM CONDITION AFTER EXTENSIVE PRACTICE. IT HAS BEEN SUGGESTED THAT THE P300 PROBABILITY EFFECT IS THE RESULT OF MEMORY UPDATING (DONCHIN, 1981). THIS SHOULD BE UNNECESSARY DURING THE AUTOMATIC MODE. THE PARADIGM EMPLOYED TO INVESTIGATE THESE ISSUES WAS SIMILAR TO THE MODIFIED STERNBERG TASK (1969) USED BY SCHNEIDER AND SHIFFRIN (1977).

PROCEDURE

Each trial began with a 10 sec presentation of the memory set. In the thrity frames that followed the presentation of each memory set, the subjects task was to press a button if a memory set item (target) was present (go-nogo task). Each of the frames contained two items, either a target and a distractor or two distractors. Each frame was presented for 200 msec. ISI's were 1600 msec.

Three variables were orthogonally manipulated in a factorial design. These variables included the number of memory set items (1 or 4), the type of practice (CM or VM) and the probability of occurence of a memory set item (.20 or .80). In the CM condition targets were always selected from one category (numbers 1 to 9) while distractors were choosen from another category (letters A to I). In the VM condition both the targets and distractors were choosen from the same category (letters A to I). Targets and distractors exchanged roles over trials in the VM conditions. Twelve sessions of 1680 trials were run with each of 5 subjects. RT's and accuracy measures were obtained in all of the sessions. ERPs were recorded in the first and twelfth sessions.

ERP RECORDING

The EEG was recorded from three midline sites (Fz, Cz & Pz) and refered to linked mastoids. Two ground electrodes were positioned on the left side of the forehead. Burden Ag-AgCl electrodes, affixed with collodion, were used for scalp and mastoid recording. Beckman bipotential electrodes, affixed with adhesive collars, were placed laterally and supra-orbitally to the right eye to record EOG and this type of electrode was also used for ground recording. Electrode impedances did not exceed 5 kohms/cm.

The EEG and EOG were amplified with Van Gogh model 50000 amplifers (time constant 10 sec and upper half amplitude of 35Hz). Both EEG and EOG were sampled for 1800 msec, begining 100 msec prior to stimulus onset. The data were digitized every 10 msec. ERP's were digitally filtered off-line (-3db at 8.8 Hz; 0 db at 20 Hz) prior to statistical analysis.

Evaluation of each EOG record for saccades and blinks was conducted off-line by calculating its variance and comparing this to a preset criterion for acceptance. Single trial EEG containing unacceptable EOG was discarded prior to statistical analysis.

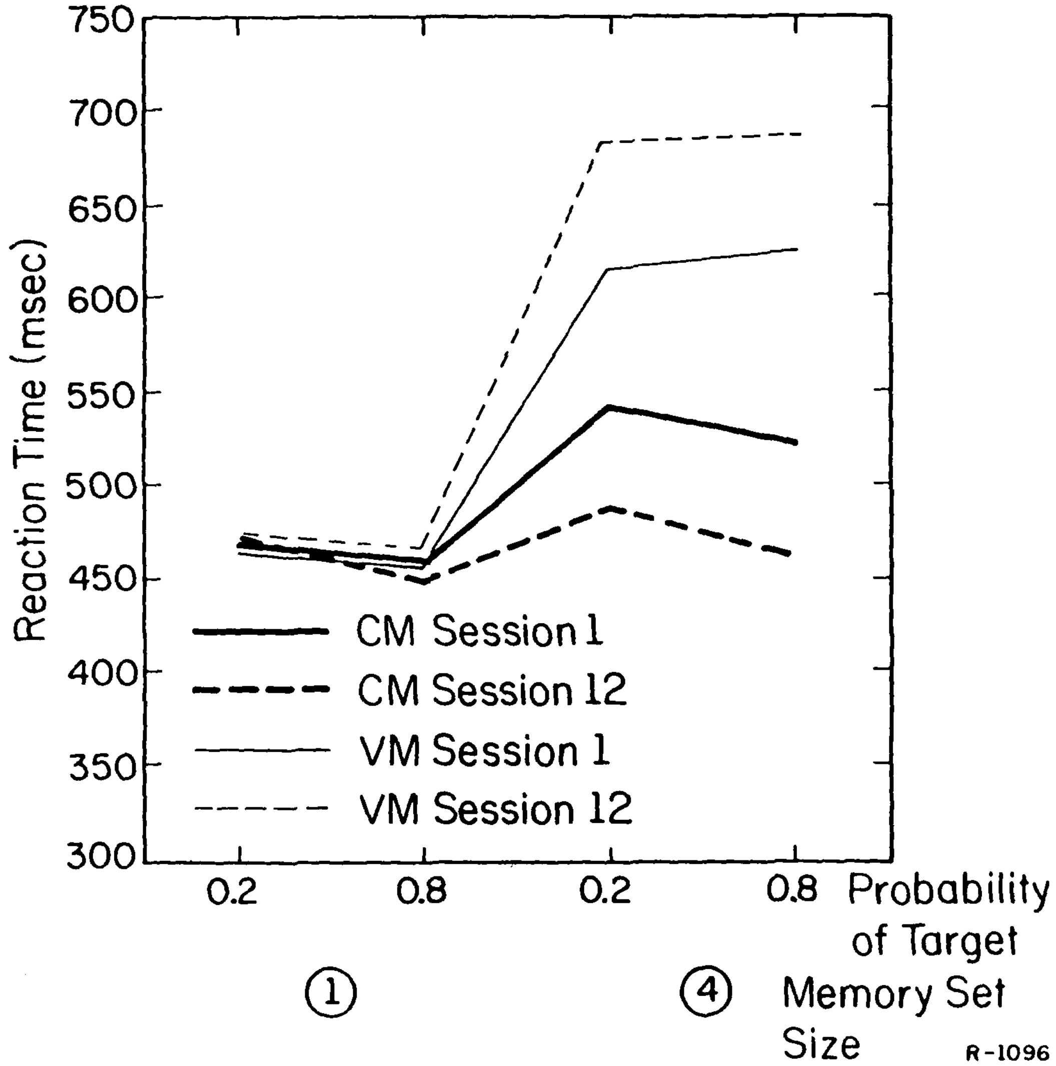

Reaction Time (msec)
750
700
650
600
550
500
450
400
350
300
CM Session 1
CM Session 12
VM Session 1
VM Session 12
0.2
0.8
0.2
0.8
Probability of Target
1
4
Memory Set Size
R-1096

Figure 1 presents the average RT's for each of the experimental conditions in sessions 1 and 12. Subjects in session 1 took longer to decide if a target was present with set size 4 than they did with set size 1. This effect was larger for VM conditions. In session 12 set size produced a significant effect for VM but not for CM conditions. This result is consistent with previous findings of a diminishing set size effect during the development of automatic processing (CM practice).

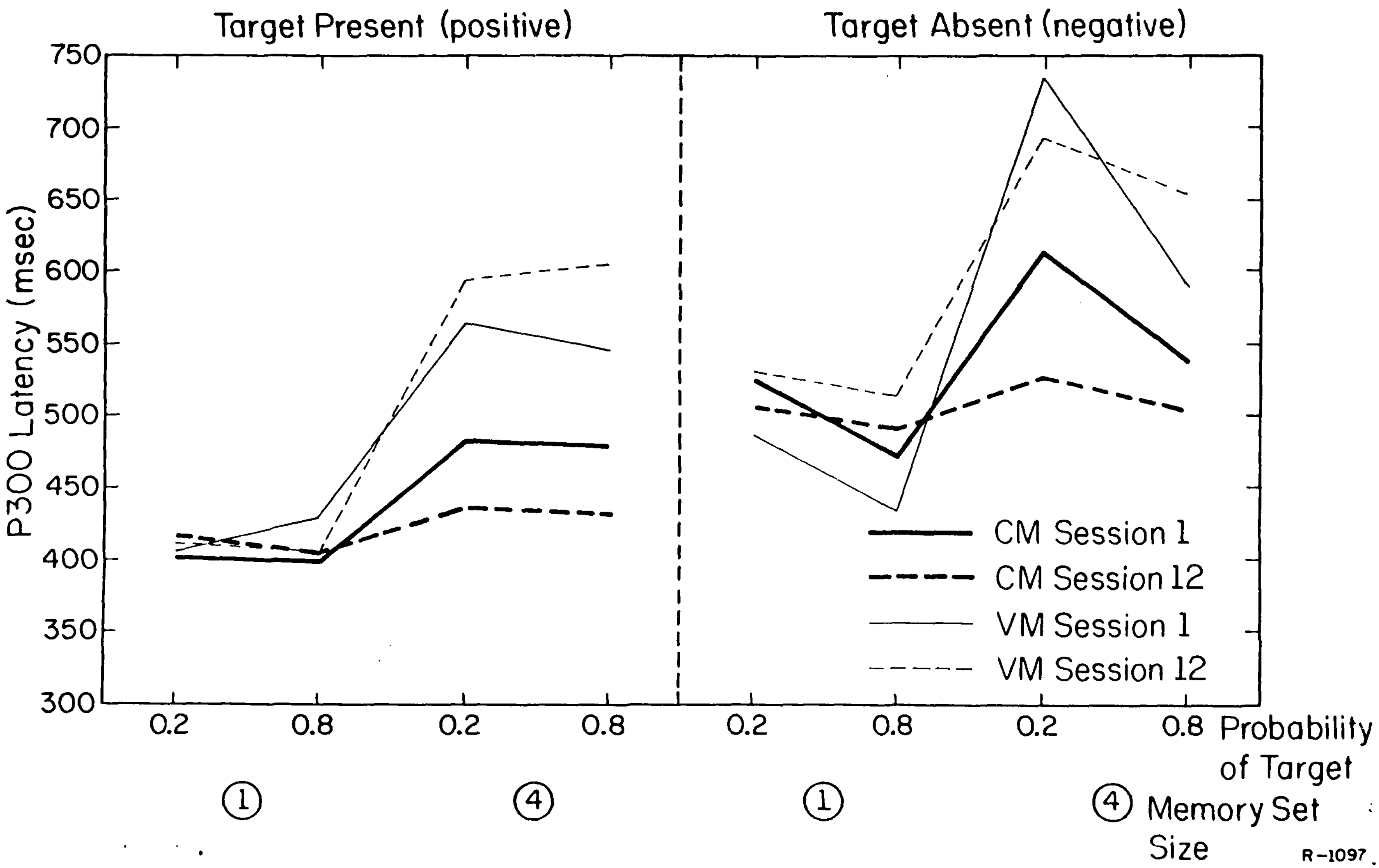
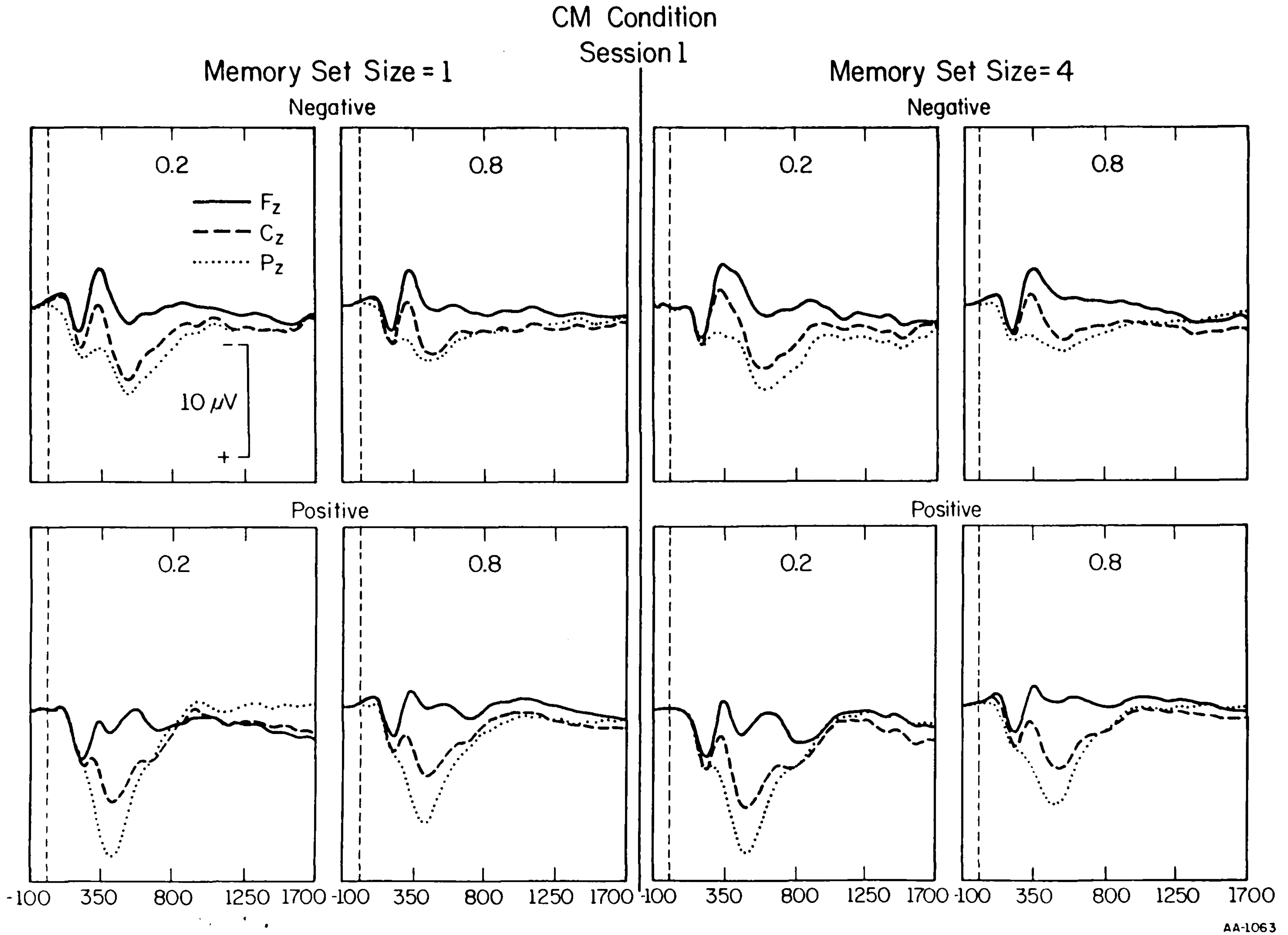
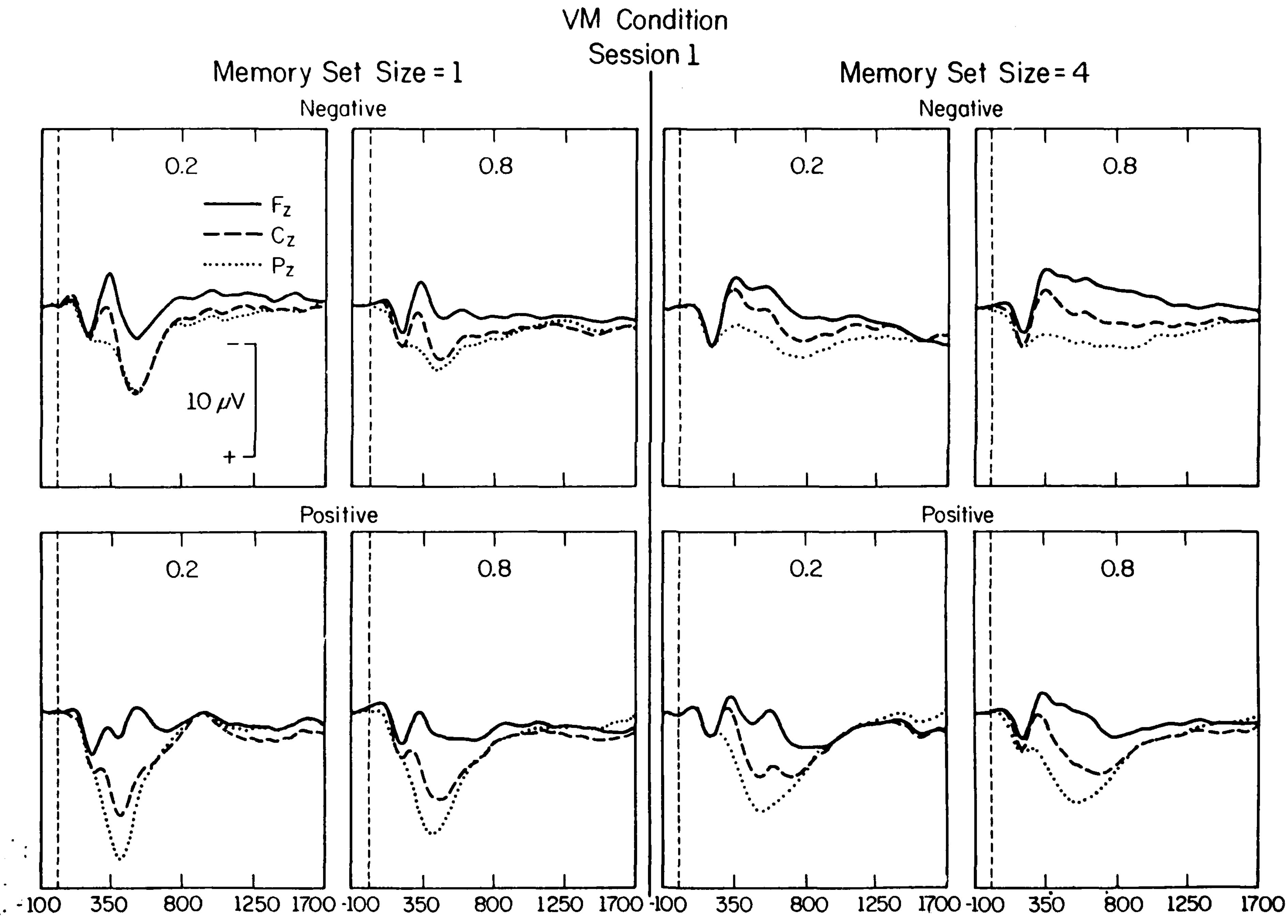
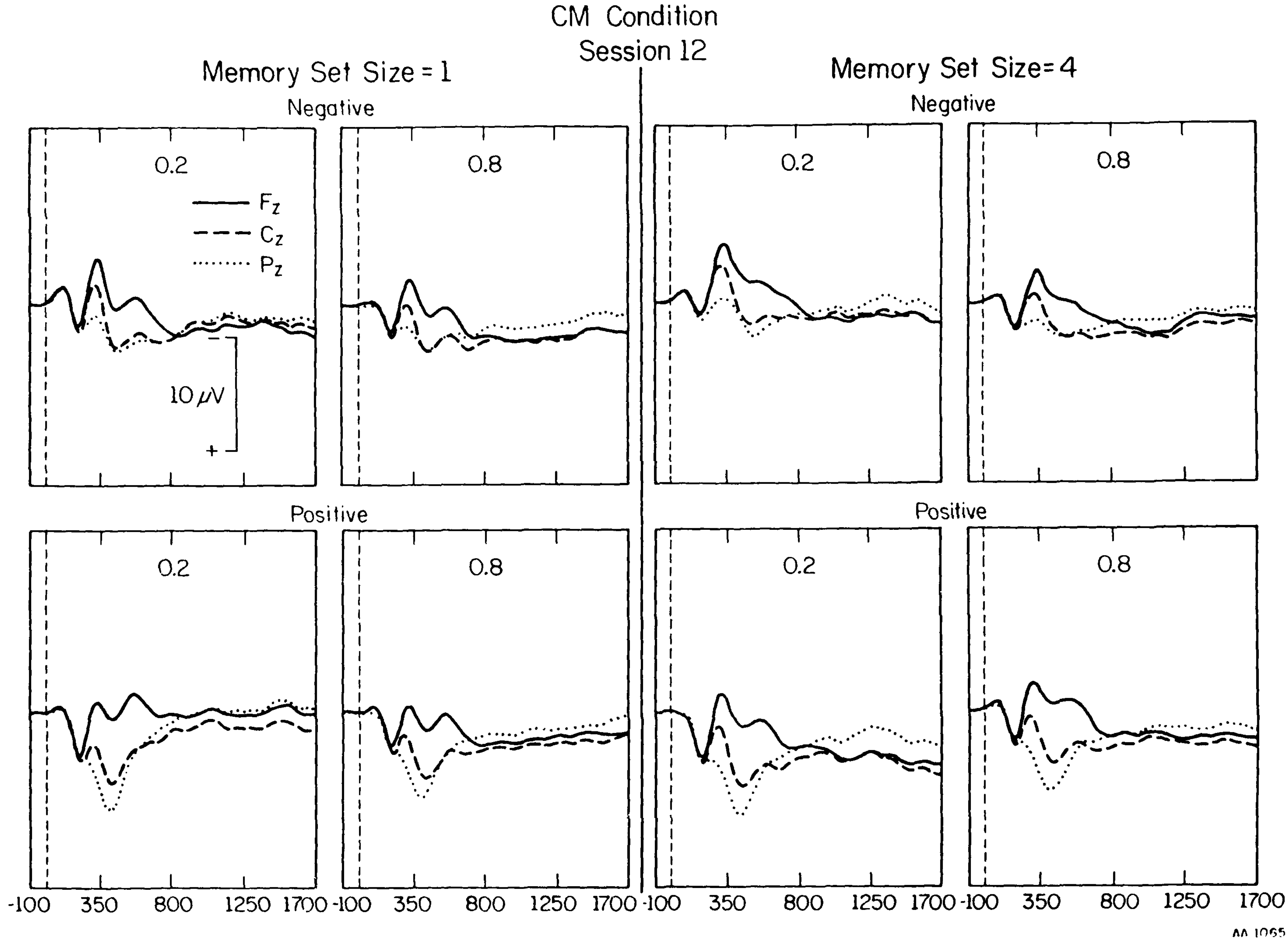

Target Present (positive)
Target Absent (negative)
P300 Latency (msec)
750
700
650
600
550
500
450
400
350
300
0.2
0.8
0.2
0.8
0.2
0.8
0.2
0.8
Probability of Target
1
4
1
4
Memory Set Size
CM Session 1
CM Session 12
VM Session 1
VM Session 12
R-1097

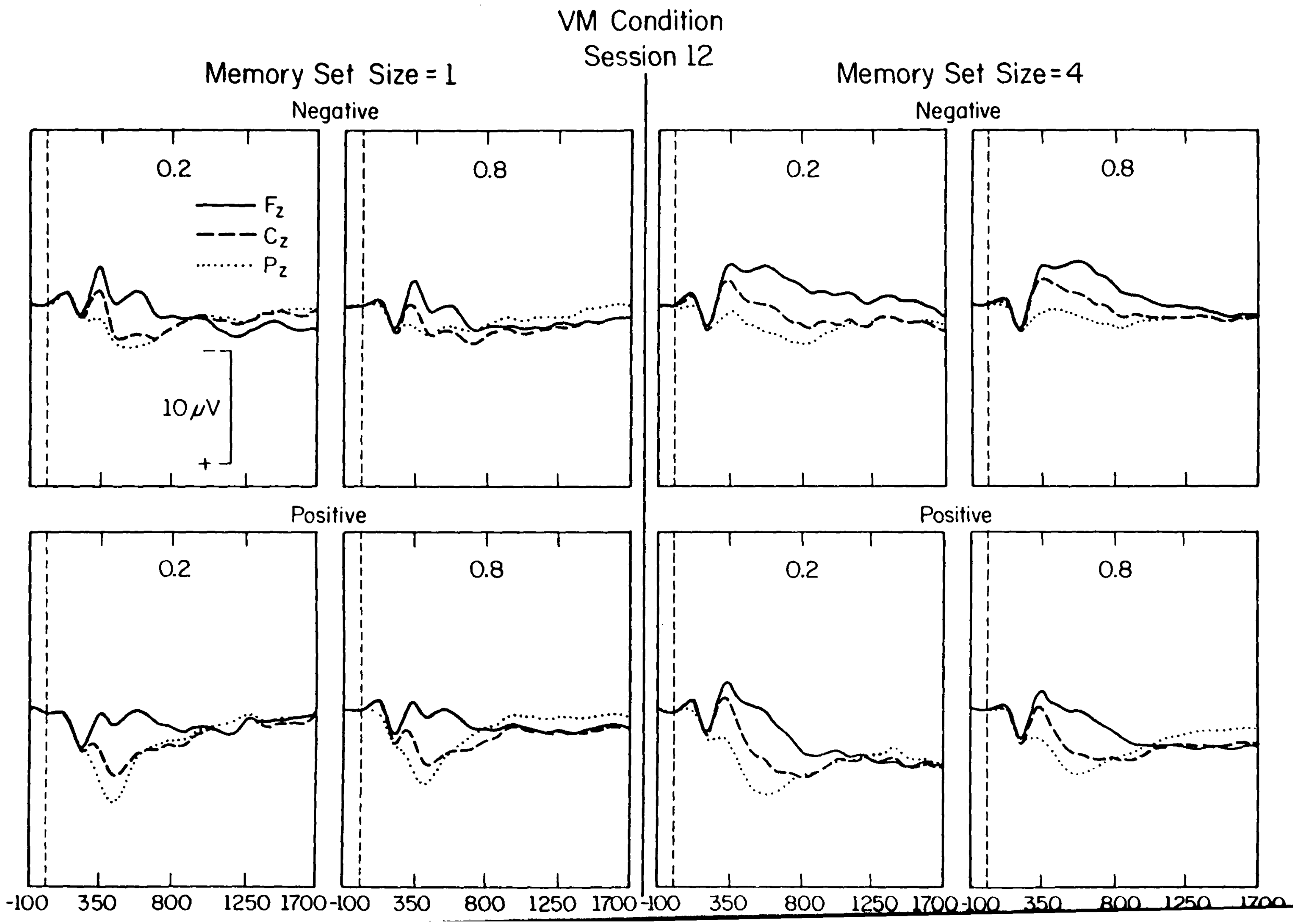
FIGURE 2 PRESENTS P300 LATENCY (OBTAINED VIA A WOODY PROCEDURE) FOR ALL EXPERIMENTAL CONDITIONS IN SESSIONS 1 AND 12. THE TYPE OF PRACTICE (CM & VM) BY SET SIZE (1 & 4) INTERACTION FOUND FOR RT WAS MIRRORED BY P300 LATENCY. SET SIZE DID NOT EFFECT P300 LATENCY AFTER EXTENSIVE CM PRACTICE. THIS FINDING SUGGESTS THAT PROCESSES OCCURRING PRIOR TO THE COMPLETION OF STIMULUS EVALUATION ARE BECOMING AUTOMATED DURING CM PRACTICE. P300 LATENCY WAS ALSO FOUND TO BE LONGER DURING THE TARGET ABSENT TRIALS THAN DURING THE TARGET PRESENT TRIALS. FURTHERMORE THE SET SIZE SLOPE FOR THE TARGET ABSENT TRIALS WAS SIGNIFICANTLY STEEPER THAN THE SLOPE FOR THE TARGET PRESENT TRIALS POSSIBLY INDICATING A SELF TERMINATING MEMORY SEARCH PROCESS.


CM Condition
Session 1
Memory Set Size = 1
Memory Set Size = 4
Negative
Positive
0.2
0.8
Fz
Cz
Pz
10 μV
-100
350
800
1250
1700
AA-1063



VM Condition
Session 1
Memory Set Size = 1
Memory Set Size = 4
Negative
Positive
0.2
0.8
Fz
Cz
Pz
10 μV
-
+
-100
350
800
1250
1700


CM Condition
Session 12
Memory Set Size = 1
Memory Set Size=4
Negative
Positive
0.2
0.8
F_z
C_z
P_z
10 μV
-
+
-100
350
800
1250
1700
AA 1065


VM Condition
Session 12
Memory Set Size = 1
Memory Set Size = 4
Negative
Positive
0.2
0.8
F_z
C_z
P_z
10 μV
+
-100
350
800
1250
1700

FIGURES 3 AND 4 REPRESENT ERP'S AVERAGED ACROSS FIVE SUBJECTS FOR EACH OF THE EXPERIMENTAL CONDITIONS IN SESSION 1. FIGURE 3 ILLUSTRATES ERP'S ELICITED DURING THE CM CONDITIONS WHILE FIGURE 4 PRESENTS ERP'S RECORDED DURING THE VM CONDITIONS. FIGURES 5 AND 6 PROVIDE THE SAME INFORMATION FOR SESSION 12. THE VERTICAL DASHED LINE REPRESENTS THE PRESENTATION OF THE STIMULUS SET.

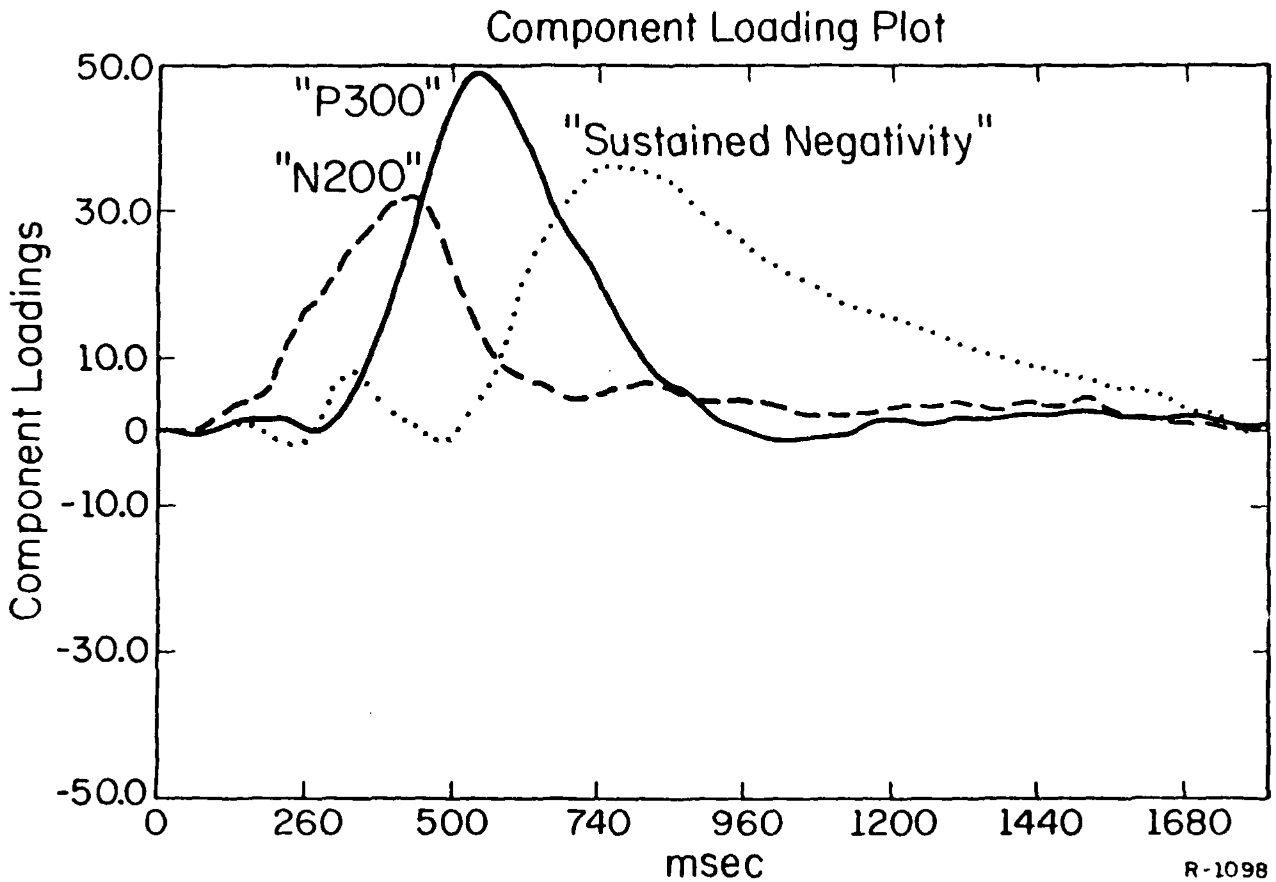

Component Loading Plot
"P300"
"N200"
"Sustained Negativity"
Component Loadings
50.0
30.0
10.0
0
-10.0
-30.0
-50.0
0
260
500
740
960
1200
1440
1680
msec
R-1098

FIGURE 7 PRESENTS THE COMPONENT LOADINGS FOR THE FIRST THREE COMPONENTS EXTRACTED FROM A PRINCIPAL COMPONENTS ANALYSIS (PCA) OF THE AVERAGE ERP'S. THE DATA BASE SUBMITTED TO THE PCA CONSISTED OF 240 ERP'S (5 SUBJECTS X 2 MEMORY SETS X 2 PROBABILITIES X 3 ELECTRODES X 2 TYPES OF PRACTICE X 2 TARGET STATES), EACH COMPOSED OF 180 TIME POINTS. THE PCA WAS PERFORMED ON THE COVARIANCE MATRIX OF THE TIME POINTS. THE THREE ASPECTS OF THE VARIANCE (COMPONENT LOADING PLOTS) WERE IDENTIFIED AS "N200", "P300" AND "SUSTAINED NEGATIVITY" BECAUSE OF THEIR TEMPORAL RELATIONSHIP TO THE STIMULUS AS WELL AS THEIR SCALP DISTRIBUTIONS.

P300 AMPLITUDE

THE AMPLITUDE OF THE P300, N200 AND SUSTAINED NEGATIVITY WERE QUANTIFIED BY THE PCA PROCEDURE. COMPONENT SCOPES DERIVED FROM THE PCA WERE SUBMITTED TO REPEATED MEASURES ANOVAS TO TEST FOR EXPERIMENTAL EFFECTS. DUE TO THE LATENCY VARIABILITY IN THE P300 COMPONENT, PCA'S WERE PERFORMED ON BOTH LATENCY ADJUSTED (WOODY PROCEDURE) AND UNADJUSTED WAVEFORMS. THE REPORTED EFFECTS ARE CONSISTENT WITH BOTH ANALYSES.

IN SESSION 1 P300'S ELICITED BY LOW PROBABILITY STIMULI WERE LARGER THAN P300'S ELICITED BY HIGH PROBABILITY STIMULI. THIS P300-PROBABILITY EFFECT IS CONSISTENT WITH PREVIOUS FINDINGS. HOWEVER, IN SESSION 12 THIS EFFECT WAS NOT FOUND IN THE CM CONDITIONS. THIS LACK OF EFFECT MAY BE DUE TO THE REDUCED NEED TO UPDATE MEMORY WHEN PERFORMING IN THE AUTOMATIC PROCESSING MODE.

P300'S ELICTED BY CM CONDITIONS WERE LARGER THAN THOSE ELICITED BY VM CONDITIONS IN SESSIONS 1 AND 12. THIS EFFECT DOES NOT APPEAR TO BE AN ARTIFACT OF INCREASED LATENCY VARIABILITY IN THE VM CONDITIONS SINCE THE DIFFERENCE REMAINED AFTER LATENCY ADJUSTMENT.

P300'S ELICITED IN THE TARGET PRESENT (POSITIVE) CONDITIONS WERE LARGER THAN P300'S RECORDED DURING THE TARGET ABSENT (NEGATIVE) CONDITIONS. ALTHOUGH PRESENCE OR ABSENCE OF THE TARGET WAS CONFOUNDED WITH THE GO-NOGO RESPONSE TASK RESULTS FROM A CHOICE RT PILOT STUDY HAVE INDICATED THAT P300 AMPLITUDE IS LARGER FOR POSITIVE TRIALS EVEN IF AN OVERT RESPONSE IS REQUIRED FOR THE NEGATIVE TRIALS. THUS THE LARGER P300 AMPLITUDE IN THE POSITIVE CONDITIONS CANNOT BE ATTRIBUTED TO THE SUPERIMPOSITION OF A MOTOR POTENTIAL ON THE P300.

N200 AMPLITUDE

THE COMMONLY OBSERVED EFFECT OF STIMULUS MISMATCH ON N200 AMPLITUDE WAS REPLICATED IN THE PRESENT STUDY. N200'S ELICITED BY TARGET ABSENT TRIALS (NEGATIVE) WERE LARGER THAN THOSE ELICITED BY TARGET PRESENT TRIALS (POSITIVE). N200'S WERE ALSO FOUND TO BE LARGER FOR MEMORY SET SIZE 4 THAN SET SIZE 1. THE LACK OF INTERACTION OF N200 EFFECTS WITH TYPE OF PRACTICE (CM OR VM) IS NOTEWORTHY. ON THE BASIS OF THE PRESENT STUDY IT APPEARS THAT THE PROCESSES REFLECTED BY N200 ARE UNAFFECTED BY THE MODE OF INFORMATION PROCESSING.

SUSTAINED NEGATIVITY

THIS FRONTALLY NEGATIVE COMPONENT OVERLAPS N200, P300 AND EXTENDS TO APPROXIMATELY 1200 MSEC POST-STIMULUS. IN SESSION 12 SUSTAINED NEGATIVITIES WERE FOUND TO BE LARGER FOR MEMORY SET SIZE 4 THAN SET SIZE 1. THIS EFFECT WAS NOT OBTAINED IN SESSION 1. THIS RESULT IS INTERESTING IN THAT TASK PRACTICE, REGARDLESS OF THE MODE OF PROCESSING (CM OR VM), INCREASES THE AMPLITUDE OF THIS COMPONENT.

IN THE VM CONDITIONS THE SUSTAINED NEGATIVITY WAS LARGER FOR THE NEGATIVE THAN THE POSITIVE TRIALS. WHEN SUBJECTS ARE OPERATING IN THE CONTROLLED PROCESSING MODE THEY MAY REQUIRE MORE PROCESSING OF THE MISMATCHES THAN WHEN THEY ARE OPERATING IN THE AUTOMATIC MODE. THE SUSTAINED NEGATIVITY MAY REFLECT THIS INCREASED MISMATCH PROCESSING.

CONCLUSIONS

THREE ENDOGENOUS COMPONENTS (N200, P300 AND SUSTAINED NEGATIVITY) USED IN CONJUNCTION WITH RT HAVE PROVIDED INSIGHTS INTO AUTOMATIC AND CONTROLLED PROCESSING MODES OF THE HUMAN INFORMATION PROCESSING SYSTEM. P300 LATENCY WAS FOUND TO MIRROR RT EFFECTS SUGGESTING THAT PROCESSES PRIOR TO THE TERMINATION OF STIMULUS EVALUATION ARE BEING AUTOMATED DURING CM PRACTICE. THE LACK OF THE P300-PROBABILITY EFFECT IN THE PRACTICED CM CONDITION POINTS TO A REDUCED NEED FOR MEMORY UPDATING DURING AUTOMATIC PROCESSING. ALTHOUGH THE N200 COMPONENT WAS EFFECTED BY MISMATCH DETECTION AND MEMORY LOAD IT WAS NOT INFLUENCED BY THE MODE OF PROCESSING. THIS SUGGESTS THAT EARLY EVALUATION OF MISMATCHES AND MEMORY COMPONENTS OF TASKS ARE PROCESSED IN THE SAME MANNER REGARDLESS OF THE LENGTH OF PRACTICE OR TASK STRUCTURE. THE SUSTAINED NEGATIVITY COMPONENT OF THE ERP PROVIDES ADDITIONAL CLARIFICATION OF MISMATCH DETECTION IN AUTOMATIC AND CONTROLLED PROCESSING MODES. SUSTAINED NEGATIVITIES WERE AFFECTED BY MISMATCHES IN THE VM BUT NOT THE CM CONDITIONS SUGGESTING THAT ADDITIONAL MISMATCH PROCESSING MAY BE NECESSARY DURING CONTROLLED PROCESSING.

Filtering for spatial distribution: A new approach (Vector Filter)

Gabriele Gratton, Michael G. H. Coles, Emanuel Donchin
Cognitive Psychophysiology Laboratory - Department of Psychology
University of Illinois at Urbana-Champaign

The values of many psychophysiological signals differ when measured at different points of the body surface. This pattern of values (where both polarity and amplitude are considered) is termed "spatial distribution". These differences reflect the distance between the source of the signal and the location of the electrodes, the nature and orientation in the space of the source, and the conductive characteristics of the interposed media. The information on spatial distribution is important for defining and describing the components of the psychophysiological signal. This is particularly true when several components contribute to the observed data, as is the case with ERPs. In this case, spatial distribution is generally considered a fundamental attribute of a component.

Vector Filtering is a statistical procedure that estimates the contribution of a particular (target) component to the data observed across several electrodes at a given timepoint. The target component is defined in terms of its spatial distribution. The estimate is based on the analysis of the similarities between the observed spatial distribution and that of the target component.

The values obtained at several electrode locations at a given time point constitute a vector (data vector). The data vector can be represented in a sp : (vector space), whose axes correspond to the electrode locations. The length of the data vector in this space is a measure of the total activity across electrode locations. The orientation of the data vector in the vector space depends on the polarity and relative amplitude at any electrode location (i.e. spatial distribution). In general, any orientation

in the vector space corresponds to one, and only one, spatial distribution. Therefore the target component can also be represented in this space as a vector (target vector), whose orientation is determined by the distribution of the target component. The similarity between the data vector and the target vector is expressed by the cosine of the angle between their orientations.

The data vector can be considered as the sum of two components. One of them is obtained by projecting the data vector on the target vector. The other is the orthogonal residual component. The degree to which this model accurately describes the data can be evaluated by testing the hypothesis that the discrepancies between a sample of data vectors and the target vector are attributable to sampling errors (Hotelling's T-square test, one-sample case). The length of the target vector is an estimate of the contribution of the target component to the observed data.

Vector Filtering has been applied to data from many experiments in which ERPs were obtained from several electrode locations in the Cognitive Psychophysiology Laboratory, including study of information processes during simple and complex tasks, memory, aging, etc. The procedure has been particularly useful in preparing data for an estimation of the latency of ERP components (P300). Some examples of the results obtained with the new procedure will be shown.

Filtering for spatial distribution: A new approach (Vector Filter)

Gabriele Gratton, Michael G. H. Coles, Emanuel Donchin
Cognitive Psychophysiology Laboratory - Department of Psychology
University of Illinois at Urbana-Champaign

INTRODUCTION

The values of many psychophysiological signals differ when measured at different points of the body surface. This pattern of values (where both polarity and amplitude are considered) is termed spatial distribution. The spatial distribution reflects the distance between the source of the signal and the location of the electrodes, the nature and orientation of the source, and the conductive characteristics of the interposed media. The information about spatial distribution is important for defining and describing the components of the psychophysiological signal. This is particularly true when several components contribute to the observed data, as is the case with ERPs. In fact, spatial distribution is generally assumed to be a fundamental attribute of an ERP component.

Vector Filtering is a statistical procedure that is based on this assumption. For any component with a prescribed distribution (target component) it provides an estimate of the amplitude of the component that is present in any observed set of data. It does this by analyzing the similarity between the observed spatial distribution and that of the target component.

PROCEDURE

Vector Filters can be used when the data set is derived from more than one electrode. In the Cognitive Psychophysiology Laboratory a minimum of three electrodes is used. However, for ease of visual representation, we consider here the two-electrode case.

Figure 1 shows average ERP waveforms obtained at two electrode sites (Cz and Pz). The values obtained across electrodes at any given timepoint (for instance, the P300 peak point) constitute a vector (data vector), that can be geometrically represented in a space (vector space), whose axes correspond to the electrode locations (Figure 2). The length of the data vector is a measure of the total activity across electrode locations. The orientation of the data vector depends on the relative amplitude at each electrode location (i.e., spatial distribution).

Any pattern of polarity and relative amplitude (i.e., any spatial distribution) corresponds to a particular orientation in the vector space. Therefore, if we believe that a component can be defined in terms of spatial distribution (for P300, Pz>Cz: see Figure 3), we can define an orientation in the vector space corresponding to that hypothetical component (Figure 4). We refer to this orientation as the target orientation .

The similarity in spatial distribution between the data vector and the hypothetical component thus defined is expressed by the angle between their orientations in the vector space. Furthermore, the projection of the data vector onto the axis defined by the target orientation provides an estimate of the amplitude of the target

component (see Figure 5). Note that the procedure of decomposing the data vector results in <u>two</u> vectors, the target vector and an orthogonal residual (<u>error</u>) vector.

Finally, we may plot the value of <u>c</u> for each timepoint in the input waveform. We term these values the <u>output</u> of the Vector Filter. Figure 6 shows the filter output for the input waveform shown in Figure 1. Note that this timeseries can be analyzed using standard data analysis procedures (e.g., peak and area measures, autocorrelation procedures, principal component analysis - PCA -, etc.).

ASSUMPTIONS OF THE PROCEDURE

1. The observed data vector can be decomposed into a target vector and a residual error vector. That is, we assume that we have chosen a target vector which represents a distribution that is actually present in the data.

2. The residual error vector does not correspond to another (target) component. That is, we assume that the target vector we have chosen is the <u>only</u> component that is present in the data.

Note: if this assumption is not met, Vector Filter can still be applied, but at least one other component with a different distribution from the target component must be invoked to explain the data.

TESTS

To test these assumptions, we derive target vector values and corresponding error vector values (c and e in Figure 5) for a particular timepoint for a set of ERP waveforms. These values are then plotted in a space defined by axes corresponding to the target and error vectors (see Figure 7). The values are used to define an ellipse corresponding to the 90% confidence region of the sample mean. Note that these values are expressed as "t scores".

Test 1: Is the target vector present in the data? This translates to the question - does the statistical distribution of the target vector values differ from zero? - which may be answered by a one-sample t-test. In this case, the 't' value was significant: hence, assumption 1 is supported. Note that, in figure 7, the ellipse does not encompass the zero value for the target vector axis. This is the visual representation of the significant difference.

Test 2: Does the residual vector contain a consistent component? Statistically the question is - does the distribution of error vector values differ from zero? In this case, the 't' value was not significant. Note that, in figure 7, the ellipse does encompass the zero value for the error vector. Thus, assumption 2 is supported.

DISCUSSION

Vector Filtering was devised to use information about spatial distribution in the analysis of psychophysiological records. We have described a procedure that allows us to test hypotheses concerning the presence or absence of specific components (defined in terms of their spatial distribution) in a certain set of data.

We can use the same kind of procedures to answer practical questions like:

- Does an observed component have a distribution similar to that generally found for P300?
- Do two groups of subjects differ in scalp distribution at a latency of 300 msec?
- For a single trial what is the latency of a component defined in terms of a particular distribution?
- What is the distribution of the component that best discriminates among two or more sets of data ?

Note that Vector Filter is able to isolate the independent contribution of several components if their orientations in the vector space are orthogonal. This is particularly important in the case of temporal overlap between components. This problem can be often overcome if appropriate choice of electrode locations is made.

FIGURE LEGENDS

Figure 1: Input of a Vector Filter. Average ERP waveforms from Cz and Pz are shown. The P300 peak point is indicated.

Figure 2: Geometrical representation of a two-element vector ($\underline{v}$). Values of the corresponding cartesian and polar coordinates are also shown. Note that two electrodes (Pz and Cz) are used as X and Y axes.

Figure 3: Scalp distribution of the target component (P300): note that Pz is more positive than Cz.

Figure 4: Orientation in the Vector Space corresponding to the target component (P300).

Figure 5: Projection of the data vector ($\underline{v}$) onto a hypothetical target component orientation. The decomposition into target vector ($\underline{c}$) and error vector ($\underline{e}$), and the corresponding values of the cartesian and polar coordinates, are also shown.

Figure 6: Output of a vector filter. The corresponding input is shown in Figure 1. The target component was P300 (see Figure 3).

Figure 7: Bivariate distribution of corresponding target and error vectors from a sample of ERPs. The ellipse indicates the 90% confidence region for the sample mean.

Fig. 1

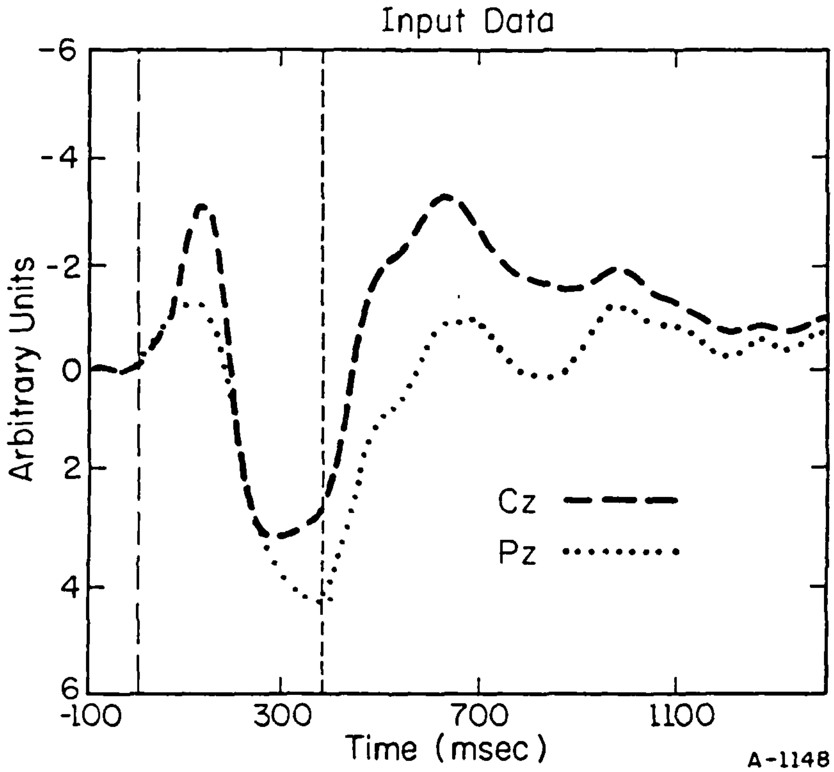
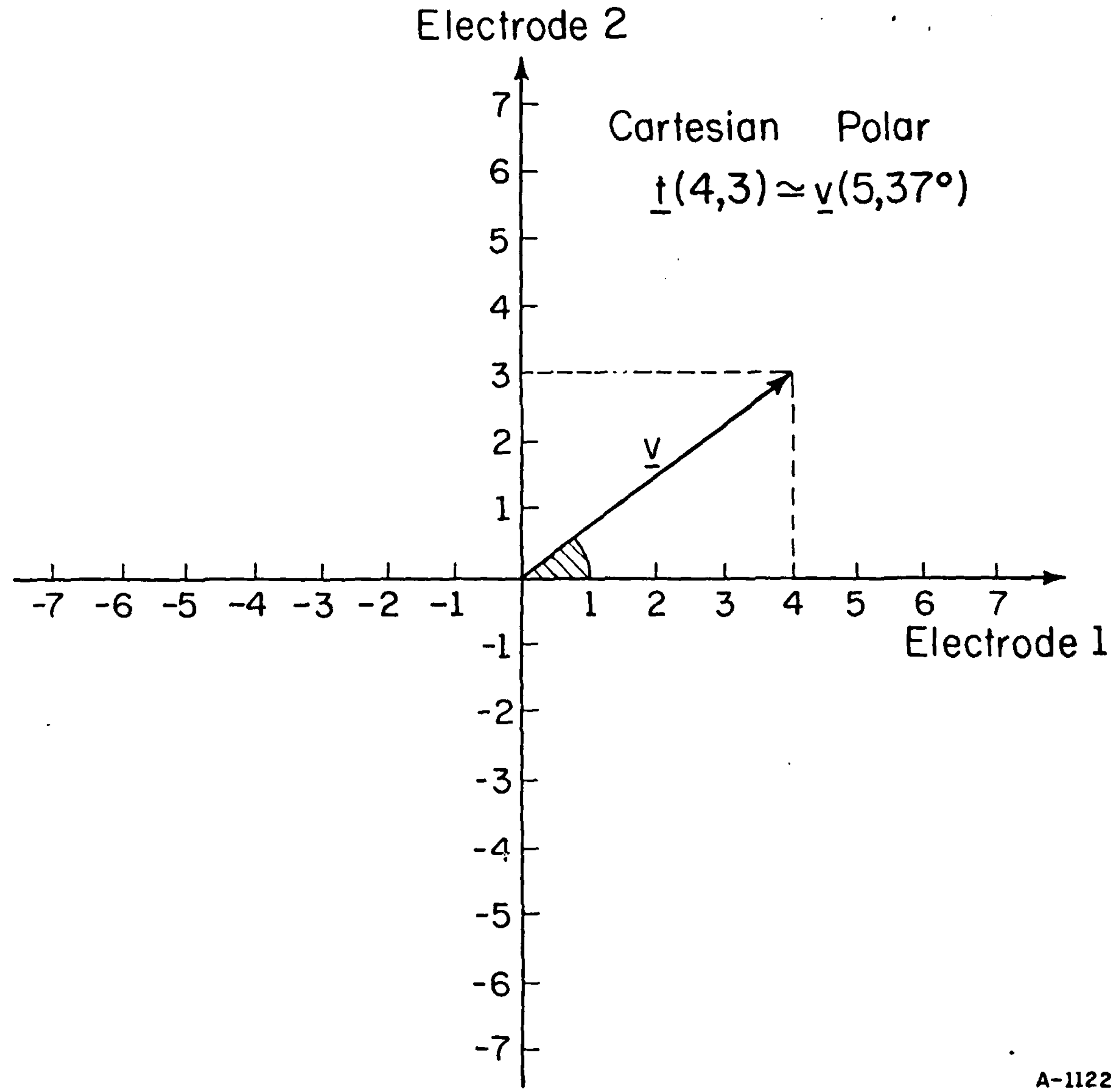

Input Data
-6
-4
-2
0
2
4
6
Arbitrary Units
-100
300
700
1100
Time (msec)
Cz
Pz
A-1148

Fig. 2



Electrode 2
Cartesian Polar
t(4,3) ≃ v(5,37°)
v
Electrode 1
-7 -6 -5 -4 -3 -2 -1 1 2 3 4 5 6 7
7 6 5 4 3 2 1 -1 -2 -3 -4 -5 -6 -7

A-1122

Fig. 3

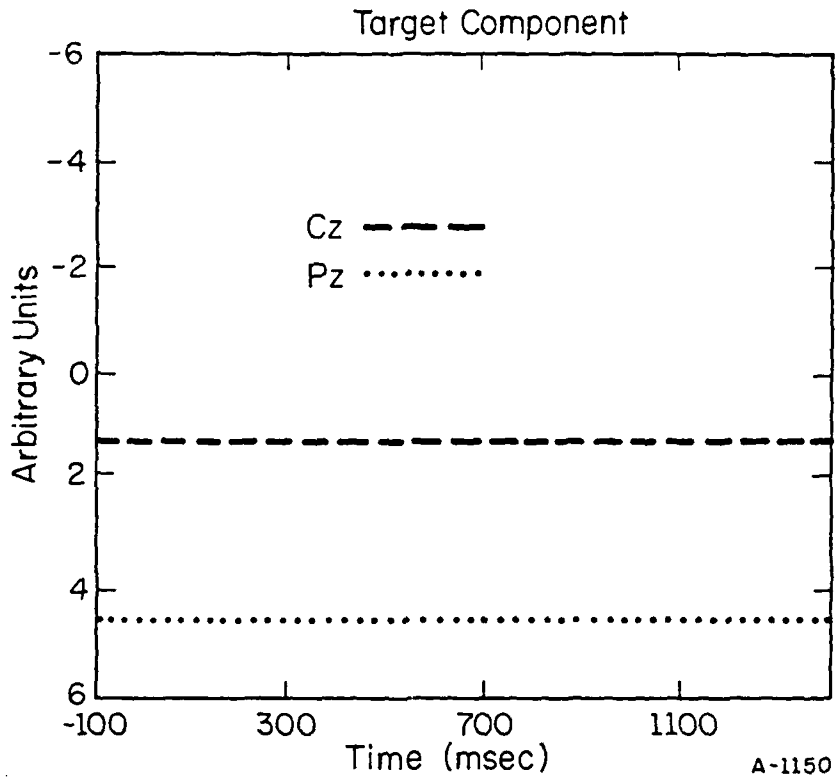
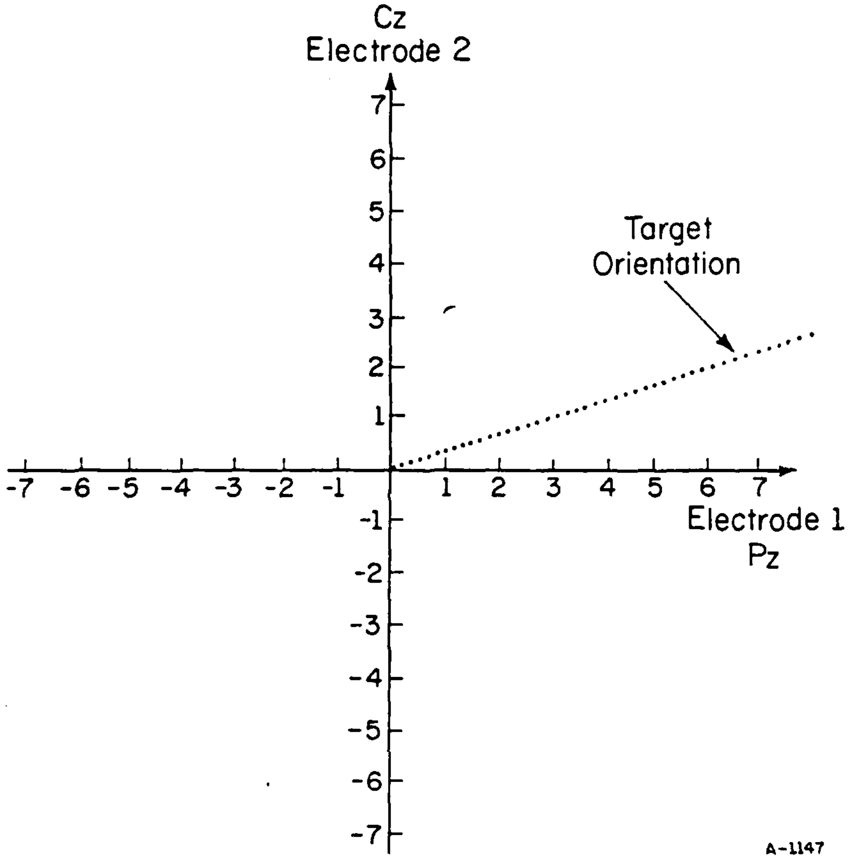

Target Component
Arbitrary Units
Cz
Pz
-6
-4
-2
0
2
4
6
-100
300
700
1100
Time (msec)
A-1150

Fig. 4



DATE 8-1
DRAWN BY K.
CHECKED BY

Fig. 5

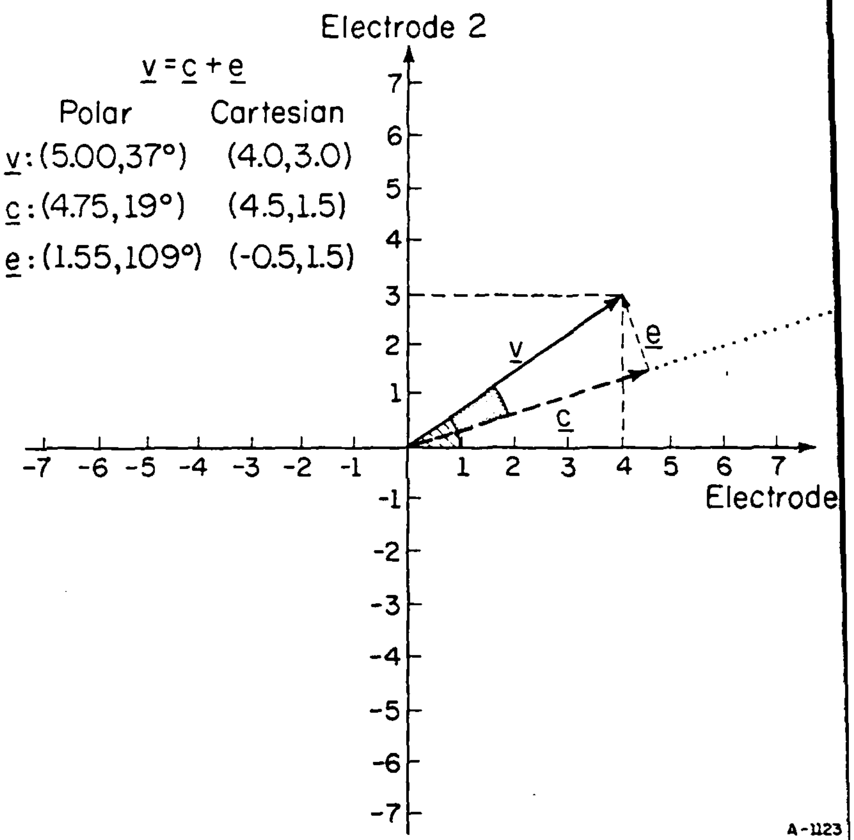

Electrode 2
v = c + e
Polar Cartesian
v: (5.00,37°) (4.0,3.0)
c: (4.75,19°) (4.5,1.5)
e: (1.55,109°) (-0.5,1.5)
v
e
c
Electrode
A-1123

Fig. 6

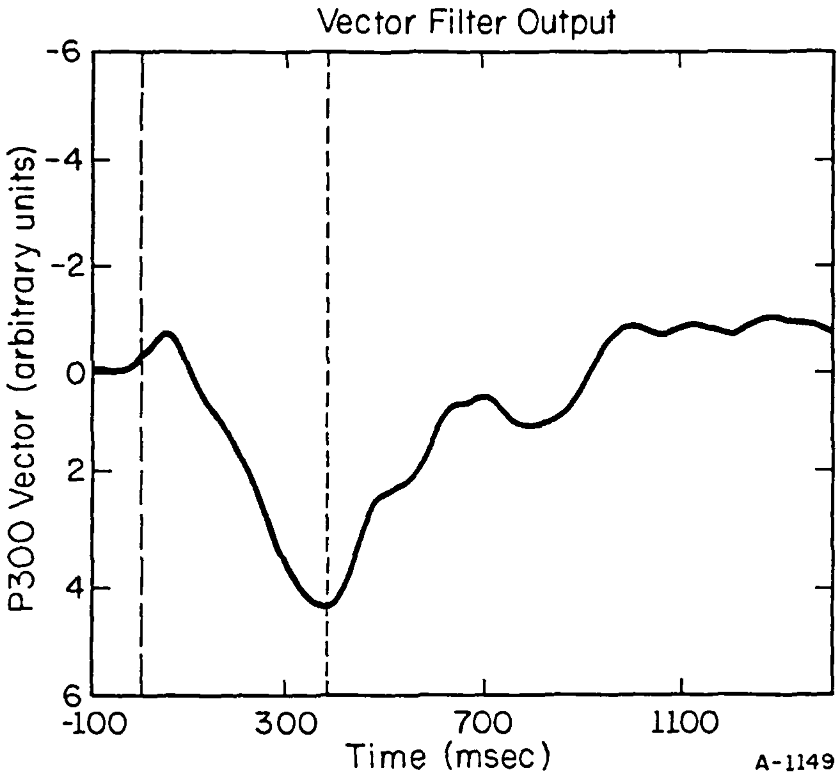
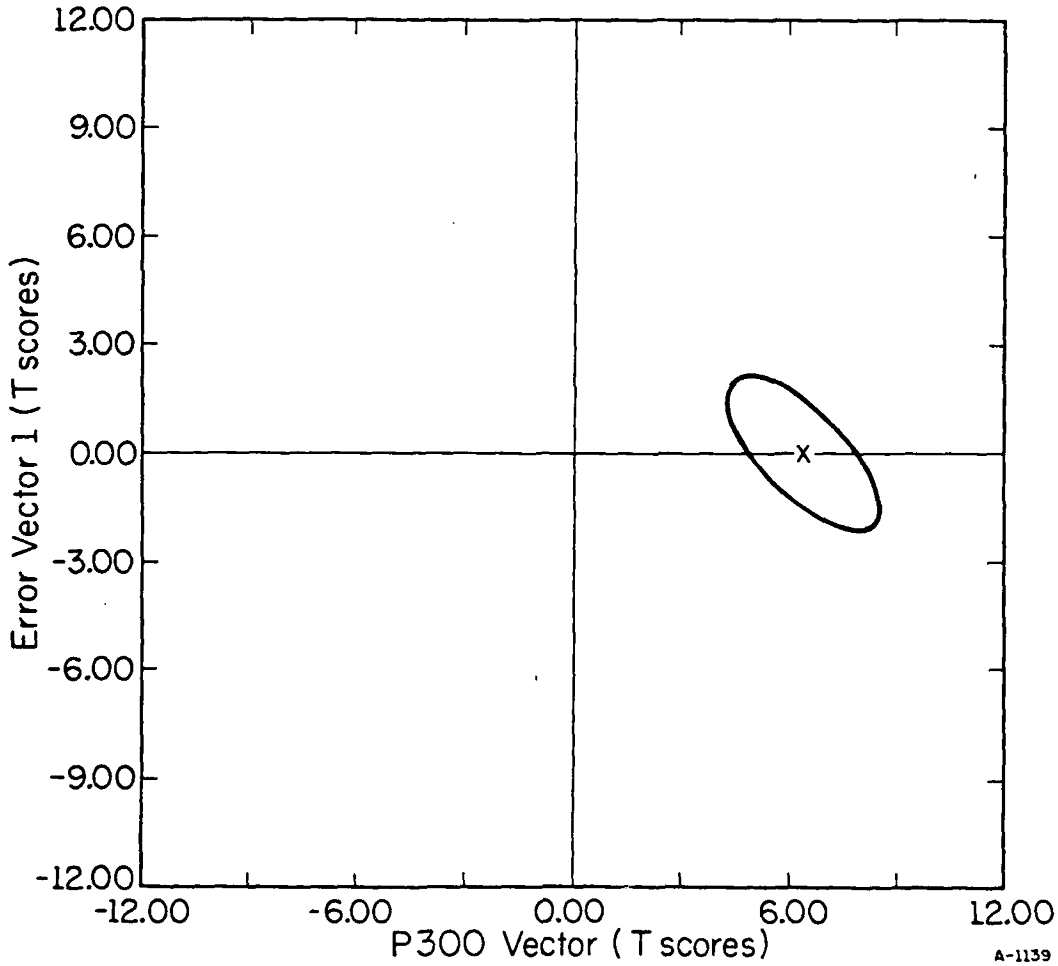

Vector Filter Output
P300 Vector (arbitrary units)
-6
-4
-2
0
2
4
6
-100
300
700
1100
Time (msec)
A-1149

Fig. 7



12.00
9.00
6.00
3.00
0.00
-3.00
-6.00
-9.00
-12.00
Error Vector 1 (T scores)
-12.00
-6.00
0.00
6.00
12.00
P300 Vector (T scores)
x
A-1139

PROCEEDINGS of the HUMAN FACTORS SOCIETY—27th ANNUAL MEETING—1983

The use of the additive factors methodology in the analysis of skill.

Amir M. Mané, Michael G.H. Coles, Christopher D. Wickens and Emanuel Donchin
Cognitive Psychophysiology Laboratory

ABSTRACT

We present an objective procedure based on the additive factors methodology for analyzing a complex task into its components. Subjects performed 16 variants of a video-game, "Space Fortress", in which four dimensions of game difficulty were manipulated orthogonally. Evaluation of the pattern of main effects and interactions for 18 performance measures revealed that the task could be broken down into two separable and one integral components. These components were associated with appraisal, motor, and perceptual-motor skills, respectively. We discuss the theoretical and practical implications of the proposed method for the design of training and for the analysis of performance deficits.

INTRODUCTION

Investigations of the performance of complex skills must often be preceeded by a decomposition of the task into an ensemble of components. The assumption underlying all such decompositions is that the task, be it the piloting of an aircraft, the control of a production plant or the writing of a book, can be viewed as a collection of sub-tasks, each challenging the operator's skills in distinct and separable ways. While there is no question that these various components display very complex interactions as they come together in the full task, there is an analytic, and often practical, convenience in examining task components separately, though this step inevitably leads to a study of the interactions between the components.

The issue arises quite clearly in the context of training procedures when one must decide whether or not training on "parts" of a task prior to full-task training is a beneficial enterprise. This question can not be answered without a prior determination of the way the task will be disassembled for the "part" training. The major controversies in this area have in fact raged around the manner in which tasks are decomposed rather than on the specific effectiveness of part-training, (Adams, 1960; Annett & Kay, 1956; Briggs & Naylor, 1962). In fact, Naylor and Briggs (1963) have argued persuasively that the effectiveness of part-training depends on the degree to which a task is decomposable.

While the importance of decomposition appears self-evident, investigators are confronted with a major hurdle. There are currently no consensual, objective, techniques for effecting such a decomposition. Much of what passes for task-analysis is essentially intuitive. The most commonly used techniques (e.g. time-line analysis) are very descriptive and can not be easily translated to a specification of the relation between the resultant components and elements of a model of the cognitive structure of the operator. That is, there is little that relates the task components to aspects of human skills and cognitive resources.

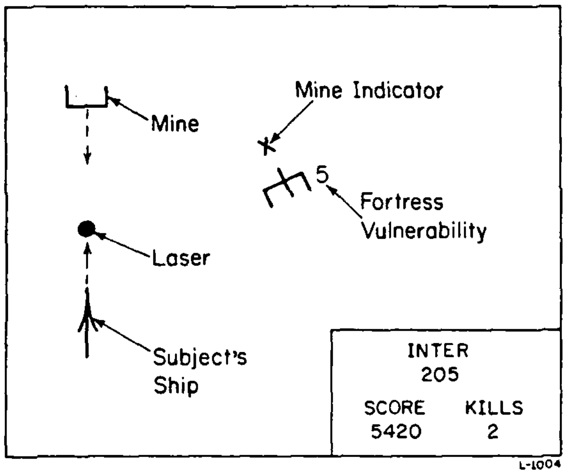
In this report we described an attempt to apply a decomposition methodology developed by Sternberg (1969) in the domain of mental chronometry to the analysis of complex tasks.Sternberg's Additive Factors approach assumes that if the effects of two independent variables are additive the two must affect independent aspects of the information processing system. Two variables whose effects on performance interact are viewed as affecting the same aspect. In the present study we have challenged subjects with a fairly complex video game. The game was so designed that its difficulty could be varied along several different dimensions. We applied these difficulty manipulations first separately, and than jointly, and studied the degree to which the effects interacted. This analysis yielded a decomposition of the task.

METHOD

The Space Fortress Task

A PDP 11-40 and an IMLAC display processor were used to present the task on a Hewlett-Packard CRT display device. Sound effects were produced by a KIM microprocessor and were presented to the subjects through a loud-speaker. The subject interacted with the display by operating a standard aviation joy-stick. The subject is seated in front of the display unit (see Figure 1) on which a number of elements are shown. His task is to destroy a Space Fortress (Fort), located in the center of the display, by pointing his Space Ship (Ship) at the Fort and firing missiles at it. In order to destroy the Fort, the subject must first hit the Fort with ten single shots, before firing a burst of two shots on target with a maximum inter-shot interval of 250 msec. The number of single hits on the Fort is displayed at all times by a digit located beside the Fort. The subject controls his Ship and fires his missiles using a joystick manipulated by his right hand. The trigger of the stick, when depressed, causes a missile to be fired by the Ship in the direction in which it is pointing. Forward movements of the stick cause

Figure 1
The Elements of the Space-Fortress Game.



the ship to accelerate in the direction in which it is pointing; lateral movements cause the ship to rotate. Because the Ship is "flying" in a frictionless environment, it will continue to fly in the direction in which it is heading unless it is rotated and thrust is applied. Thus, control of the Ship is a complex perceptual-motor task.

In trying to destroy the Fort, the subject is thwarted by a number of different obstacles. First, the Fort can rotate, "lock-on", and fire missiles at the subject's Ship. Thus, the subject cannot remain stationary. Second, from time to time, a mine emerges from the Fort and chases the subject's Ship. Every missile that is shot when a mine is present on the screen is ineffective against the Fort. The mines can be of two types, friend or foe, and the subject must act differently depending on the mine type. Mines are identified by a character from the alphabet which appears above the Fort when a mine emerges. The subject is told before any given run of the task which characters indicate foe mines. If the mine is a foe, the subject must first identify it as such before firing a missile to destroy it. Identification is accomplished by the depression of a button located on top of the joystick. The subject must depress this button twice, with a prescribed inter-press interval, to accomplish identification. If the mine is a friend, no identification response is required, and the mine can be "energized" by a single missile shot. If the subject fails to destroy a foe mine or to energize a friend mine within 10 sec, the mine will self destruct. The interval between mine appearances is 3 sec, and the subject must use this interval to fire at the Fort.

For the purposes of training, and for the task analysis procedure, the difficulty of the task was varied along four dimensions: (a) Memory set size (either 1 or 5): the number of characters that could identify a foe mine was either 1 or 5. The set was given to the subject before each run. (b) Mine speed (either fast or slow): the speed of either friend or foe mines was either 15 or 30 units of display speed. The maximum speed of the subject's ship was 40. (c) Identification response interval (easy or hard): the interval between button presses used to identify foe mines was either 100-350 or 250-450. (d) Mine blinking (on or off): in the "on" condition, the mines (friend or foe) could disappear for 1500 msec, although they remained on the screen for at least 1500 msec after emerging. This cycling of 1500 msec on/1500 msec off continued until either the mines were destroyed or energized or the subject's ship was destroyed. While invisible, the mines continued to pursue their normal course and could destroy the Ship or be destroyed. Note that, for each manipulation, their are two levels. Thus, the combination of every level of every variable with every other level yields 2 x 2 x 2 x 2, or 16, different conditions, with each condition defined by a particular level of each of the four variables.

Subjects

Five subjects, recruited from the university community, were paid $3.50 per hour for participating in the experiment. These subjects had received between 30 and 60 hours of training on simple versions of the task. During training, subjects were presented with an easy version of the task, in which all task dimensions were at their easiest levels. Then, as subjects achieved mastery at the easy level of the task, the difficulty level was increased until subjects showed mastery at the most difficult level. The criteria for mastery were: less than two Ship kills and more than 10 Fort kills in two consecutive five minutes runs.

Procedure

Each of the 16 conditions was performed in two consecutive blocks, yielding 32 five min blocks per subject, distributed over four sessions. For each subject, the order of conditions within and across sessions was determined according to Latin square procedures. A 5 min "warm-up" period preceded each session.

Performance measures

On every run, 18 measures of the subjects performance were derived. The measures included: score which is a composite index of performance, ships killed by mine, by fort, and total; Fort hits, kills, and average time to kill a fort; Mine kills, % foe mines not destroyed, % friend

mines not energized, and efficiency of mine shots. The process of identifying and killing a mine yielded the following measures: time to identification of foe, time to kill a foe, time from ID to kill (the difference between the last two measures), time to energize friend, extra time to kill a foe, (the difference between the last measure and time to kill a foe). Accuracy was measured by number of bad identifications (friend identified as foe), wasted shots (shots at foe identified as friend) and number of bad intervals (interval was outside the specified range).

RESULTS AND DISCUSSION

To evaluate the pattern of main effects and interactions relating the manipulation of memory set, mine speed, identification response interval, and blinking, to the performance measures, we first employed a 2 x 2 x 2 x 2 x 2 x 5 analysis of variance. The last two factors correspond to the replication and subjects factors, respectively. The results of this analysis revealed clearly that the mine speed variable interacted with other variables with respect to many of the performance measures. To obtain a clearer picture of the underlying structure of the task, we decided to perform two separate analyses, one for each of the two levels of mine speed.

Before we turn to the interpretation of these data, we should note that, in the first analysis, with mine speed included as a factor, five performance measures were significantly influenced by mine speed. For these same five measures, no interactions between mine speed and the other variables were evident. Table 1 presents the main effects of the four independent variables on 5 of the primary performance measures. Score, number of fort hits, and time to destroy fort, indicate superior performance with increased mine speed. However, increasing mine speed led to an increase in the number of ships killed, this effect being due to the number of ships killed by a mine rather than by the Fort. These latter two performance measures (ships killed and ships killed by mine) were also influenced by the three other manipulations. In all cases, as difficulty increased, so more Ships were killed.

The results of the analyses for slow and fast mine speed conditions separately are shown in Table 2. The blinking manipulation is not included in the the table because none of the performance measures was significantly influenced by this manipulation. In Table 2, means surrounded by parentheses were not significantly different ($p<.05$), although they were in the expected direction ($p<.20$). For each measure, the upper and lower rows give the results for the analysis for slow and fast mine speeds, respectively.

	MS		Mem		Int		Bl	
General Measures	Slow	Fast	1	5	Easy	Hard	Off	On
score	1904	2013	-	-	-	-	-	-
ships killed	1.2	4.8	2.7	3.4	2.5	3.5	2.6	3.4
Fortress Measures								
number of fort hits	203	221	-	-	-	-	-	-
time to destroy fort	17.2	15.1	-	-	-	-	-	-
Mine Measures								
ships killed by mine	0.8	4.6	2.3	3.1	2.2	3.2	2.3	3.1

Table 1

Means for Significant Main Effects from Analysis of Variance on Performance Measures with Mine Speed (MS), Memory Set Size (Mem), ID Response Interval (Int), and Blinking (Bl) as Factors.

Mine Measures	Mem		Int	
	1	5	Easy	Hard
% foe mines not destroyed	-	-	-	-
	15.7	18.7	14.7	19.7
% friend mines not energized	-	-	-	-
	(7.3	8.3)	-	-
number of bad IDs	-	-	-	-
	.17	.53	-	-
% of wasted shots	-	-	-	-
	4.2	7.4	-	-
time to ID foe	-	-	-	-
	.89	.99	-	-
time to energize friend	-	-	-	-
	1.49	1.58	-	-
ID to kill time foe	-	-	1.32	1.52
	-	-	.82	1.04
time to kill foe	-	-	-	-
	-	-	1.79	1.97
extra time to kill foe	-	-	-	-
	-	-	.26	.43
efficiency of mine shots	-	-	-	-
	-	-	68.7	72.2
bad intervals	-	-	-	-
	-	-	.63	1.83

Table 2

Means for Significant Main Effects from Analysis of Variance on Performance Measures for the two Mine Speed Conditions Separately. For Each Performance Measure, the Upper and Lower Rows give Means for Slow and Fast Speeds Respectively.

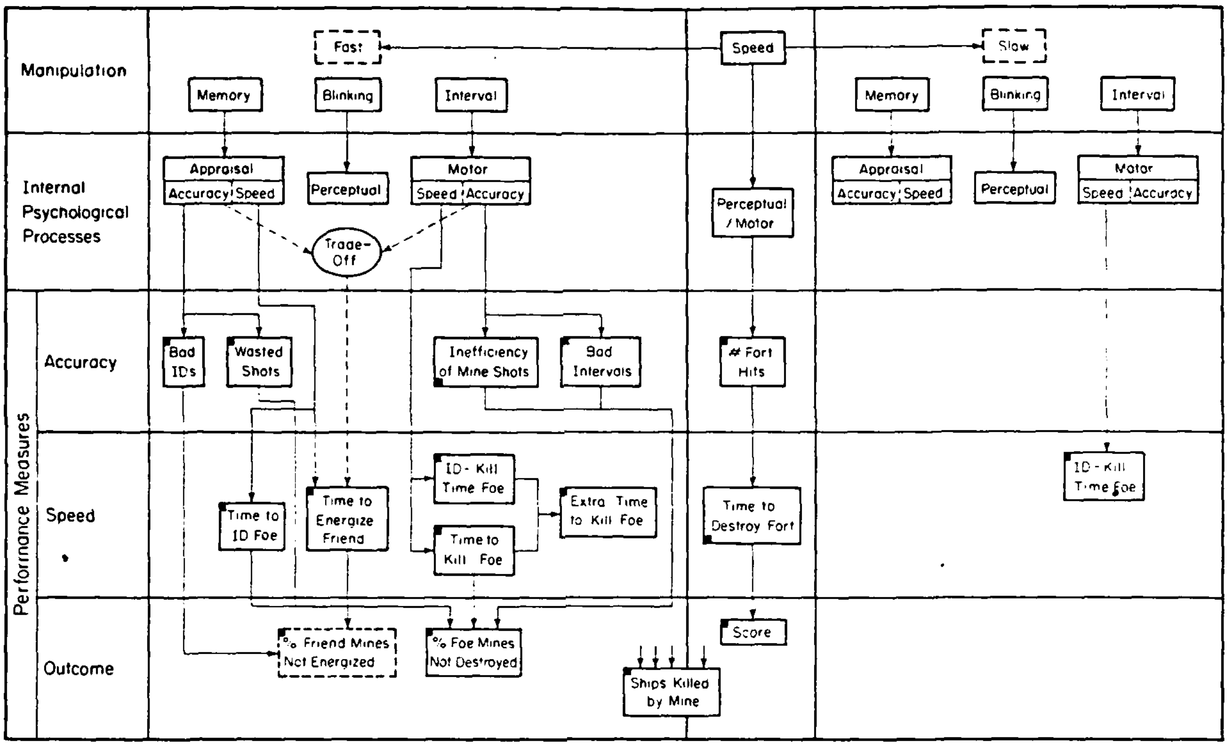


Figure 2
The Task Structure, Inferred from the Results of the Analysis of Variance.

Scrutiny of the pattern of results shown in Table 2 reveals that there are essentially two classes of performance measures - those influenced by memory set size and those influenced by the ID interval response requirement. This immediately suggests that the two classes of measures are tapping into two separable aspects of the subject's performance which correspond in turn to independent skills at performing different components of the tasks. Furthermore, for each of the two classes, there appears to be a speed and accuracy aspect.

Figure 2 shows the structure of the task as inferred from the pattern of results shown in Tables 1 and 2. Note first that for the slow mine speed condition (right-hand panel), only one performance measure is influenced by the other manipulations. This suggests that, in handling the slow speed, the subjects had spare capacity available to cope with the increased demands implied by increases in memory and motor requirements. However, the requirements inherent in coping with the fast mine speed appear to lead to a drain on some central processing resources such that the other variables now exert a profound influence on many of the performance measures (left-hand panel). It is interesting to note that, although more ships are killed under the fast mine speed, other aspects of the subject's performance actually improve. This is particularly the case with those measures related to destruction of the Fort. Thus, the subject hits the Fort more frequently and destroys the Fort more quickly when the mine speed is fast. We infer, therefore, that the increase in allocation of resources to those perceptual-motor processes involved in dealing with the increase in mine speed (presumably relating to general aspects of Ship manipulation) results in a generalized improvement in those behaviors which depend on perceptual-motor processes (see center panel).

The data for the fast mine speed are easy to interpret if we consider, in detail, the several stages of action that must occur between the appearance of a mine and its destruction. First, the subject must identify if the mine is a friend or a foe. If it is a friend, he may proceed to energize it by flying his Ship into the appropriate location and then firing his missile. If the mine is a foe, then he must first produce the appropriate identification response (double button press) before proceeding to destroy it. As with friend mines, this is accomplished by flying the ship into the appropriate location and then firing a missile. The pattern of results given in Table 2 and displayed in Figure 2 conform to this analysis. First, we note that memory set size has two major influences. It has an effect on both accuracy of iden[illegible]ication (identifying a foe as a friend - bad I[illegible] or vice versa - wasted shots) and on the speed of the identification

process (time to ID foe and time to energize friend). Second, if the mine is a foe, we note that the interval requirement exerts an influence. Again, this influence is manifested in both speed and accuracy measures. For speed, increasing the difficulty of the identification response leads to an increase in the time between the identification of a foe and its destruction and, correspondingly, in the overall time to kill a foe and the extra time to kill a foe. For accuracy, increasing the difficulty of the identification response leads to an increase in the number of incorrect identifiction intervals and, more generally, in the number of shots at the mine which hit their target (efficiency of mine shots). In turn, variations in the demands placed on the appraisal and motor processes lead to variations in the outcome measures, percent of foe mines destroyed and, although not significantly, on the percent of friend mines energized.

Of all the possible interactions between the memory and interval manipulations, only one is significant, that for time to energize a friend. Scrutiny of this interaction suggests a trade-off between appraisal and motor processes. When both memory set size and interval are most difficult, performance is most degraded.

These results indicate that the task can be analyzed into at least three components: appraisal, motor, and perceptual-motor. The appraisal and motor components are essentially isolable. The perceptual-motor component is not isolable since it interacts with the other two components. The lack of influence of the "blinking" manipulation on performance measures is informative. On the one hand, we cannot determine what skill component is related to this manipulation, on the other hand, it is clear that no special training is necessary for mastery of the skill.

CONCLUSIONS

The results of the present study suggest the following conclusions:

1. The analysis of complex tasks can be aided by the additive factors methodlogy. In our case, three skill components underlie performance of the Space Fortress task.
2. Note that we have related the components to those psychological resources and stages of processing that are proposed by current theories in cognitive/human engineering psychology (e.g Wickens, 1980).
3. Our results suggest that the skills associated with memory and motor aspects of the task might be acquired through part training. However, for a subject to achieve proficiency, training of the task in its entirety (including the perceptual-motor aspect) must be given. Design of training strategies, and evaluation of their effectiveness, are topics for future investigation. However, the present research provides clear guidelines and predictions.

ACKNOWLEDGEMENTS

This research was supported by contract # F49620-79-C-0233 from the Air Force Office of Scientific research, by contracts # F33615-82-C-0609 and # F33615-82-K-0611 from Brooks AFB, and by contract # MDA903-83-C-0017 from Defense Advanced Research Projects Agency.

REFERENCES

Adams, J. A. Part trainers. In G. Finch (Ed.) Educational and Training Media: A symposium. Washington, D. C.: National Academy of Sciences - National Research Council 1960.

Annett, J., & Kay, H. Skilled performance. Occupational Psychology, 1956, 30, 112-117.

Briggs G. E., & Naylor, S. C. The relative efficiency of several training methods as a function of transfer task complexity. Journal of Experimental Psychology, 1962, 64, 505-512.

Naylor, J. C., & Briggs, G. E. Effects of task complexity and task organization on the relative efficiency of part and whole training methods. Journal of Experimental Psychology, 1963, 65, 217-224.

Stammers, P. Part and whole practice for a tracking task: effects of task variables and amount of practice. Perceptual and Motor Skills,1980, 50, 203-210.

Sternberg, S. On the discovery of processing stages: Some extensions of Donders' method. Acta Psychologica,1969,30,276-315.

Wickens, C. D. The structure of attentional resources. In R. Nickerson (Ed.), Attention and Performance, VIII, Englewood Cliffs, NJ: Erlbaum, 1980.

ERPs and Performance under Stress Conditions

A. Mane, E. Sirevaag, M. G. H. Coles, & E. Donchin
Cognitive Psychophysiology Laboratory
University of Illinois

The present study explored the utility of measures of the event-related potential (ERP) in the analysis of complex task performance under stress. Five expert and five novice subjects played a video game, "Space Fortress", for 12 consecutive hours. The "mission" was performed once during the day and again at night. The subject controlled an armed space ship in a hostile environment. For most of the mission, a low level "vigilance" task was presented. At intervals which averaged 30 sec, an element on the display (space fortress) flashed. A bright flash (p=.2) indicated that six sec later a mine would accelerate and pursue the ship. The subject had to take immediate action to avoid the mine and destroy it. A dim flash (p=.8) had no consequences.

P300 and CNV components were evident in the ERPs, recorded from Fz, Cz, and Pz, following the fortress flashes. Both components were larger after the bright (low probability/target) flashes. For the experts, the amplitude of both components decreased over time. For both experts and novices, CNV amplitude was smaller during the night mission, while P300 amplitude was larger at night for experts only. Variation in the amplitude of both components was related to variation

in different aspects of the subjects' performance. These results indicate (a) that measures of the ERP are consistent even over a 12 hour recording period, (b) that changes do occur over time. There is some indication that these changes are associated with different types of performance decrement.

ERPs and Performance under Stress Conditions

A. Mane, E. Sirevaag, M. G. H. Coles, & E. Donchin
Cognitive Psychophysiology Laboratory
University of Illinois at Urbana-Champaign

INTRODUCTION

The present study explored the utility of measures of the event-related potential (ERP) in the analysis of performance of a complex task under the stress of a twelve hour session. The task was a complex video game, "Space Fortress" which was developed for research purposes. In a previous experiment (Mane, Coles, Wickens, & Donchin 1983) we used the additive factors methodology to analyze the task into components (see Figure 1). The questions addressed in the present research were:

1. Can "traditional" ERP components be recorded in a highly complex task environment and over an extended period of time?
2. How do time, shift (day vs. night), and expertise relate to these components?
3. Is there an association between ERP measures and performance measures? Can such an association be understood in terms of the task structure?

PROCEDURE

Five expert and five novice subjects participated in two 12 hour sessions, one during the day (8 am - 8 pm) and one at night (8 pm - 8am). The long duration mission consisted of two states: Vigilance and Hell (see Figure 2). Performance measures are shown in Figure 1.

Vigilance task (see Figure 2)

An enemy mine slowly pursued a space ship which was controlled by the subject. At intervals which averaged 30 sec, an element on the display (space fortress) flashed. A bright flash (p=.2) indicated that six sec later a mine would accelerate and pursue the ship. The subject had to take immediate action to avoid the mine and destroy it. A dim flash (p=.8) had no consequences.

Hell (see videotape)

Six times during the 12 hour mission, a highly- demanding version of the game was presented for 5 mins. The subject controlled an armed space ship, equipped with lasers and missiles. The ship performed in an hostile environment which consisted of a stationary fortress capable of shooting at the ship and space mines (either friends or foes) which pursued the ship and destroyed it upon contact. The object of the game was to activate friendly mines, to shoot foe mines and to destroy the fortress. Subjects were paid both a flat rate and bonus for good performance. Prior to the experiment the experts received an average of 60 hours of training and the novices 16 hours.

ERP RECORDING

The EEG was recorded from three midline sites (Fz, Cz and Pz according to the 10-20 system) and referred to linked mastoids. Burden Ag-AgCl electrodes were used for scalp and mastoid and recording. Beckman Biopotential electrodes, affixed with adhesive collars, were placed laterally and supra-orbitally to the right eye to record EOG, and this type of electrode was also used for ground. Electrode impedance did not exceed 10 Kohms/cm. The EEG and EOG were amplified with Van Gogh model 50000 amplifiers (time constant 10 sec and upper half amplitude of 35 Hz). Both EEG and EOG were sampled for 1280 msec, beginning 100 msec prior to the stimulus onset. ERPs were recorded following both dim and bright fortress flashs (see Figure 2). The data were digitized every 10 msec. ERPs were digitally filtered off-line (-3db at 8.3 Hz; 0db at 20 Hz) prior to statistical analysis. Eye movement artifacts were corrected off line using a procedure described by Gratton Coles and Donchin (1983). Reported results are based on analyses using PCA and the vector filter procedure described by Gratton Coles and Donchin (Science Fair 1).

RESULTS

A. ERP data

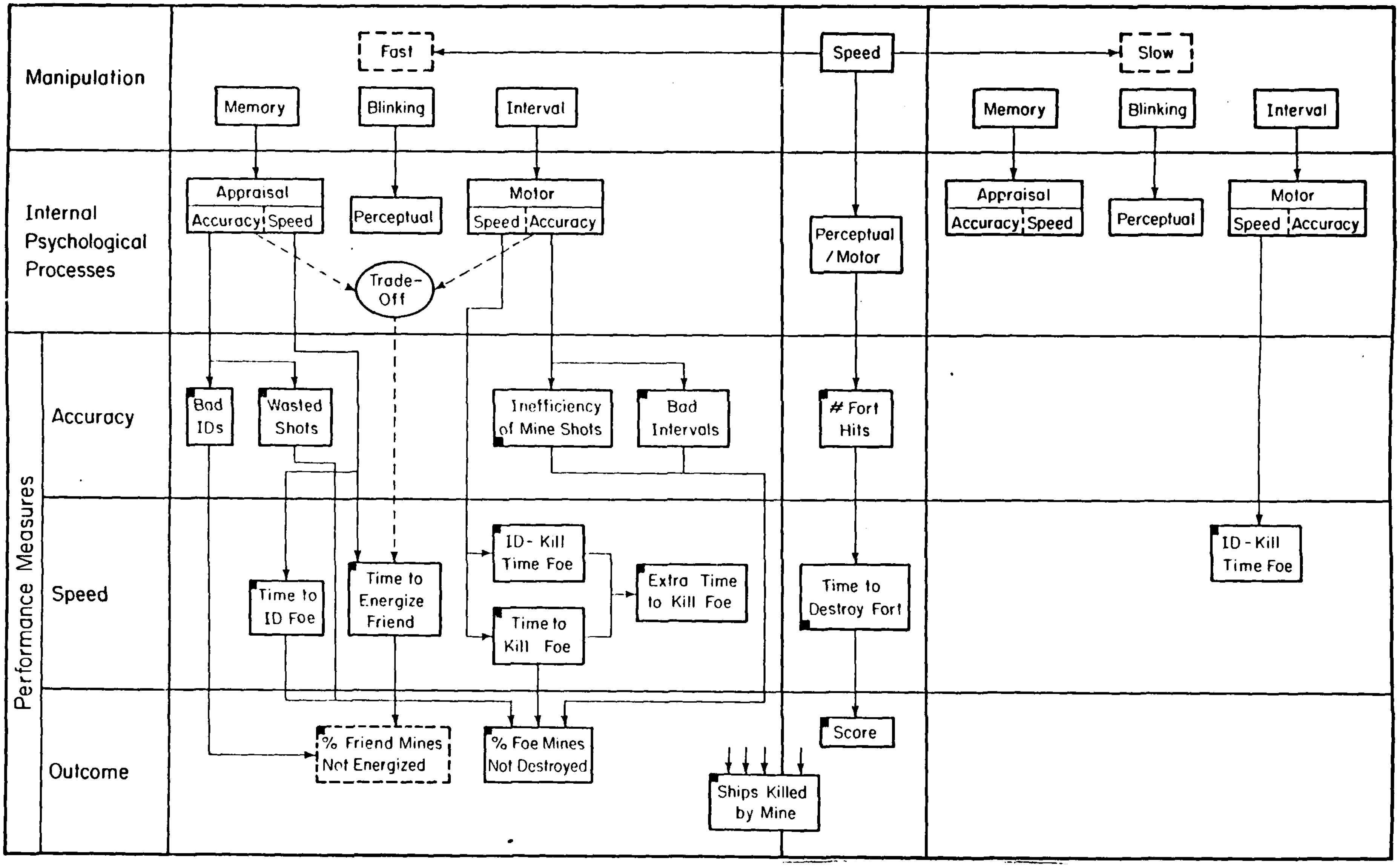
1. The bright, less probable (20%), target flashes elicited both a P300 and a CNV. No CNV and a small P300 were elicited by the dim flashes (see Figure 3).
2. For the experts, the amplitude of both components decreased over time. For both experts and novices, CNV amplitude was smaller during the night mission, while P300 amplitude was larger at night for experts only (see Figure 3).
3. Although a difference between novices and experts is apparent in the waveforms representing the response to the dim flashes (see Figure 4), traditional methods of ERP analysis have failed to reveal a statistically significant difference between groups.

B. ERP-performance relationships

To evaluate these relationships, each subject's ERP data for the 12 vigilance only blocks were identified. Then, for a particular ERP measure (e.g. P300 amplitude), the two blocks containing the highest and lowest values for that measure were isolated. The significance of the differences in performance measures between these two blocks was then assessed. Results are shown in Figure 5a.

This procedure was repeated for the Hell performance data from the blocks adjacent to the vigilance blocks from which the ERP data had been derived. Results are shown in Figure 5b.

Note that, in both cases, ERP-performance relationships are present.



L-1015

Figure 1. Task structure based on the additive factors analysis.

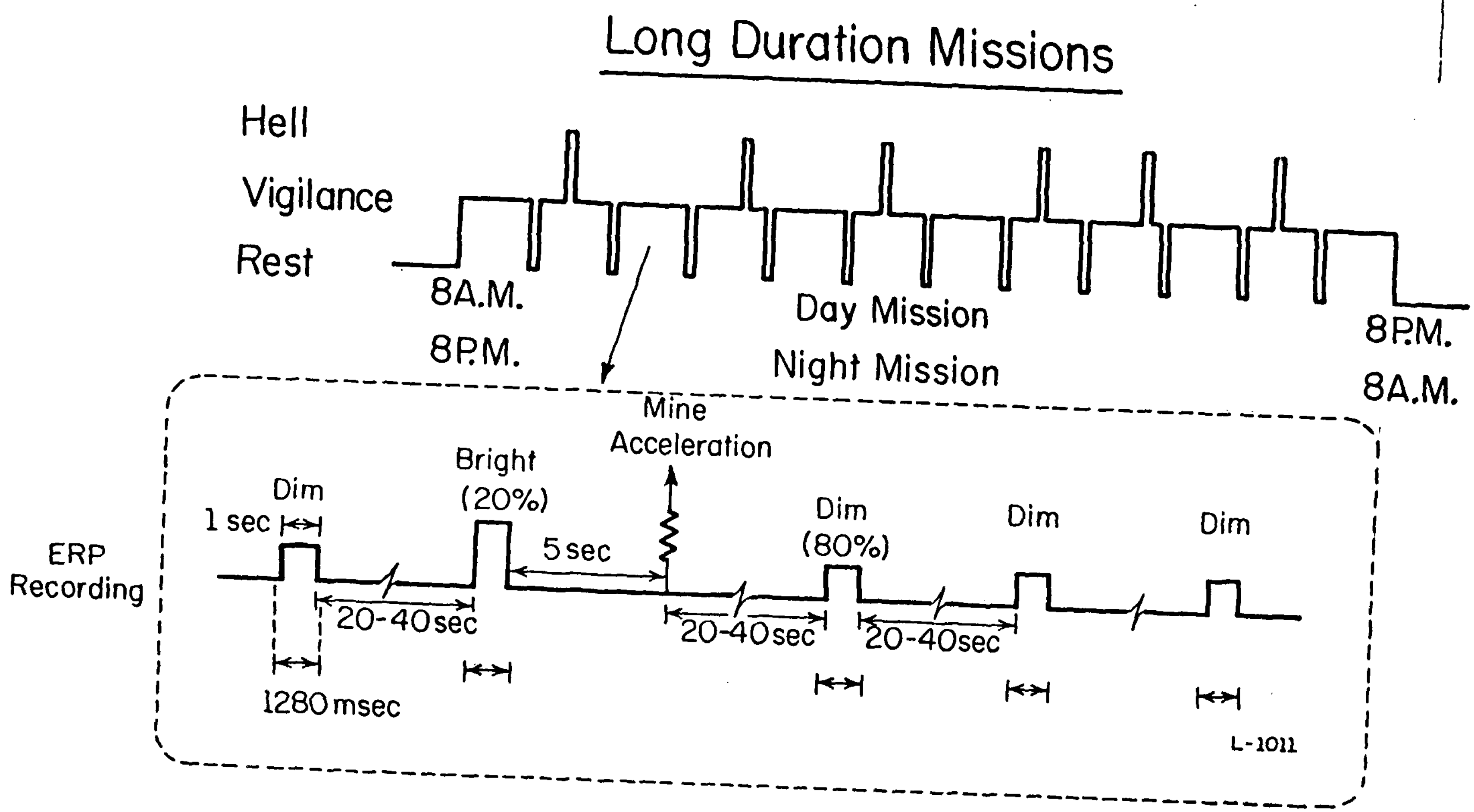
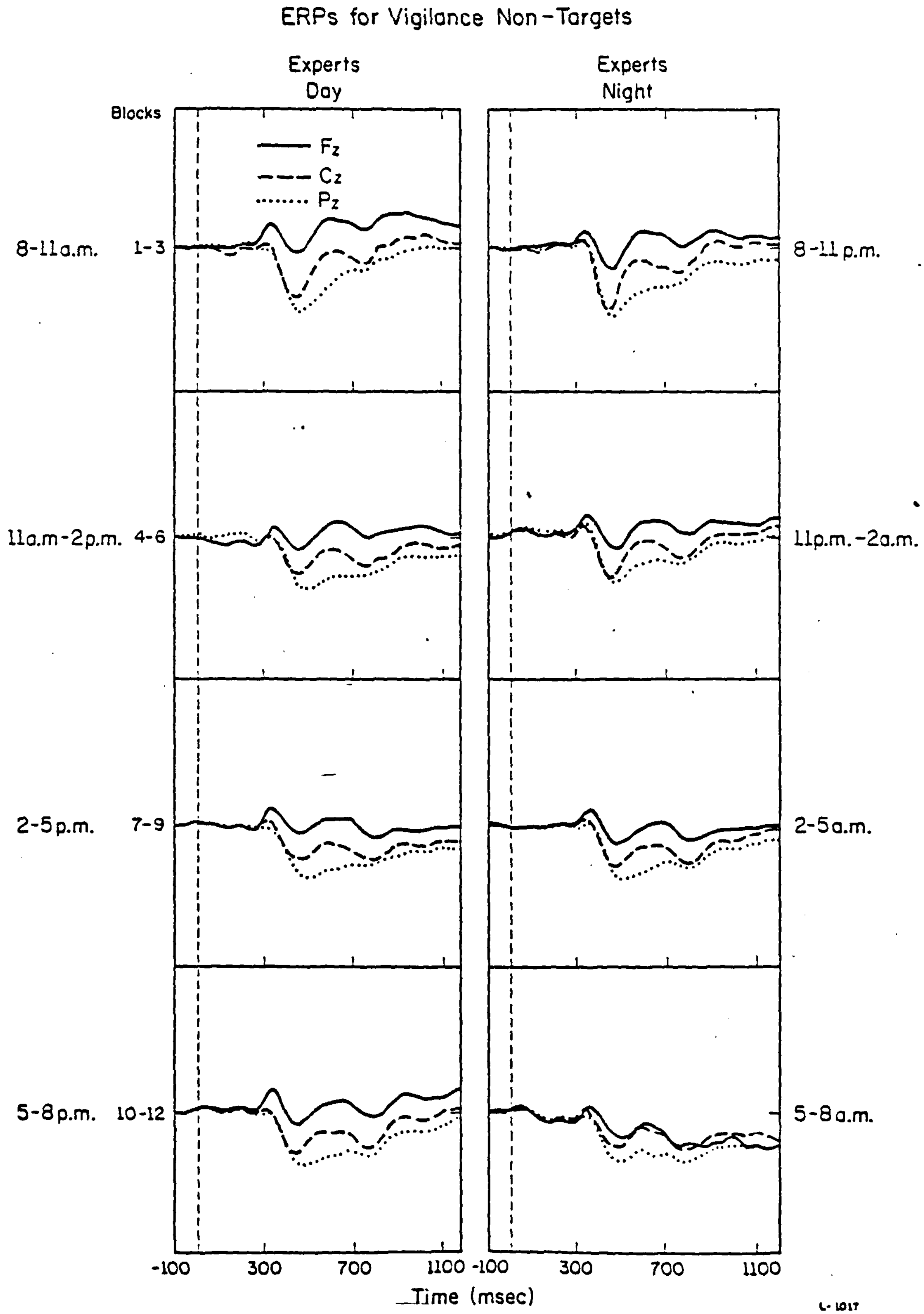
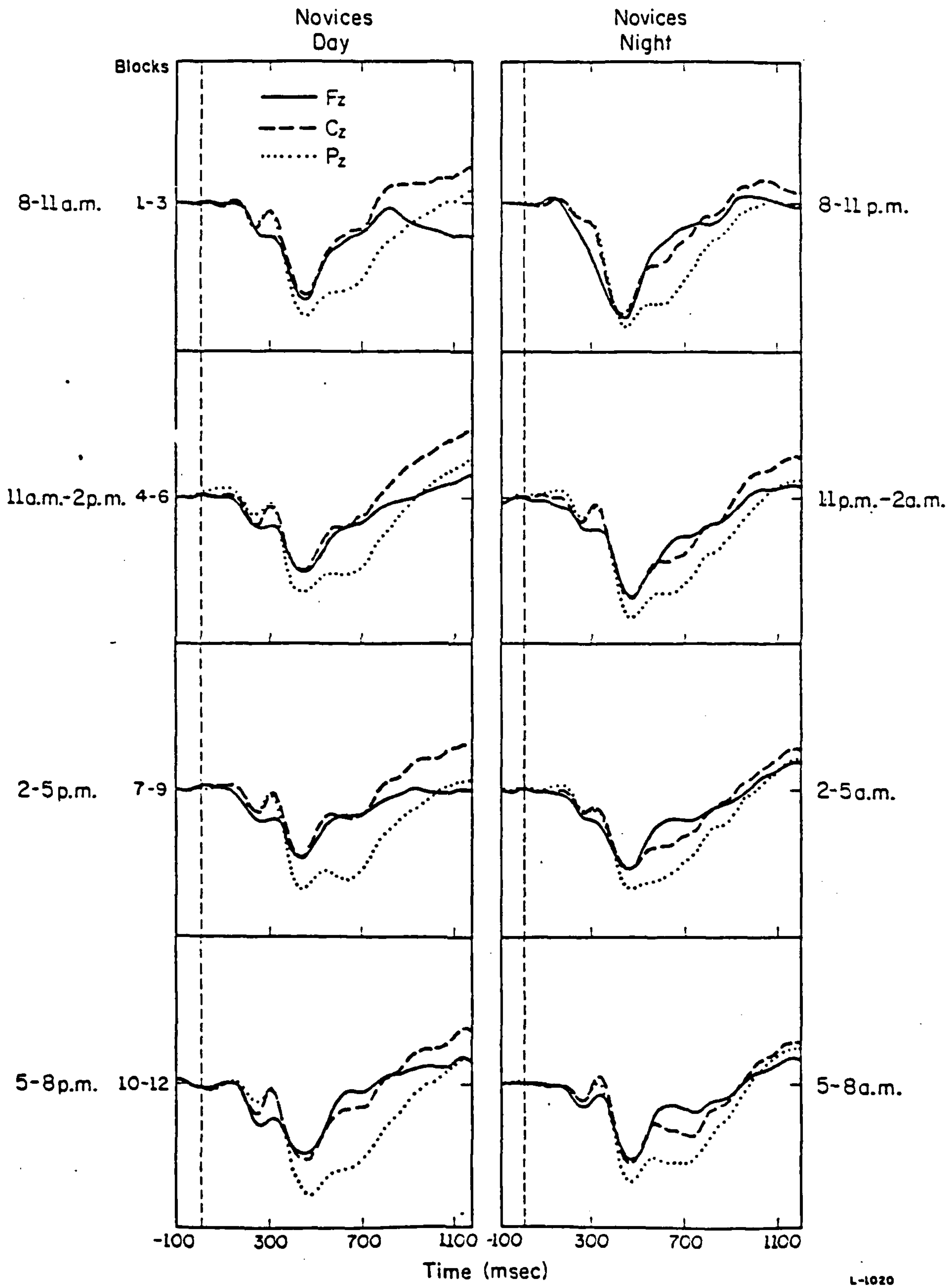


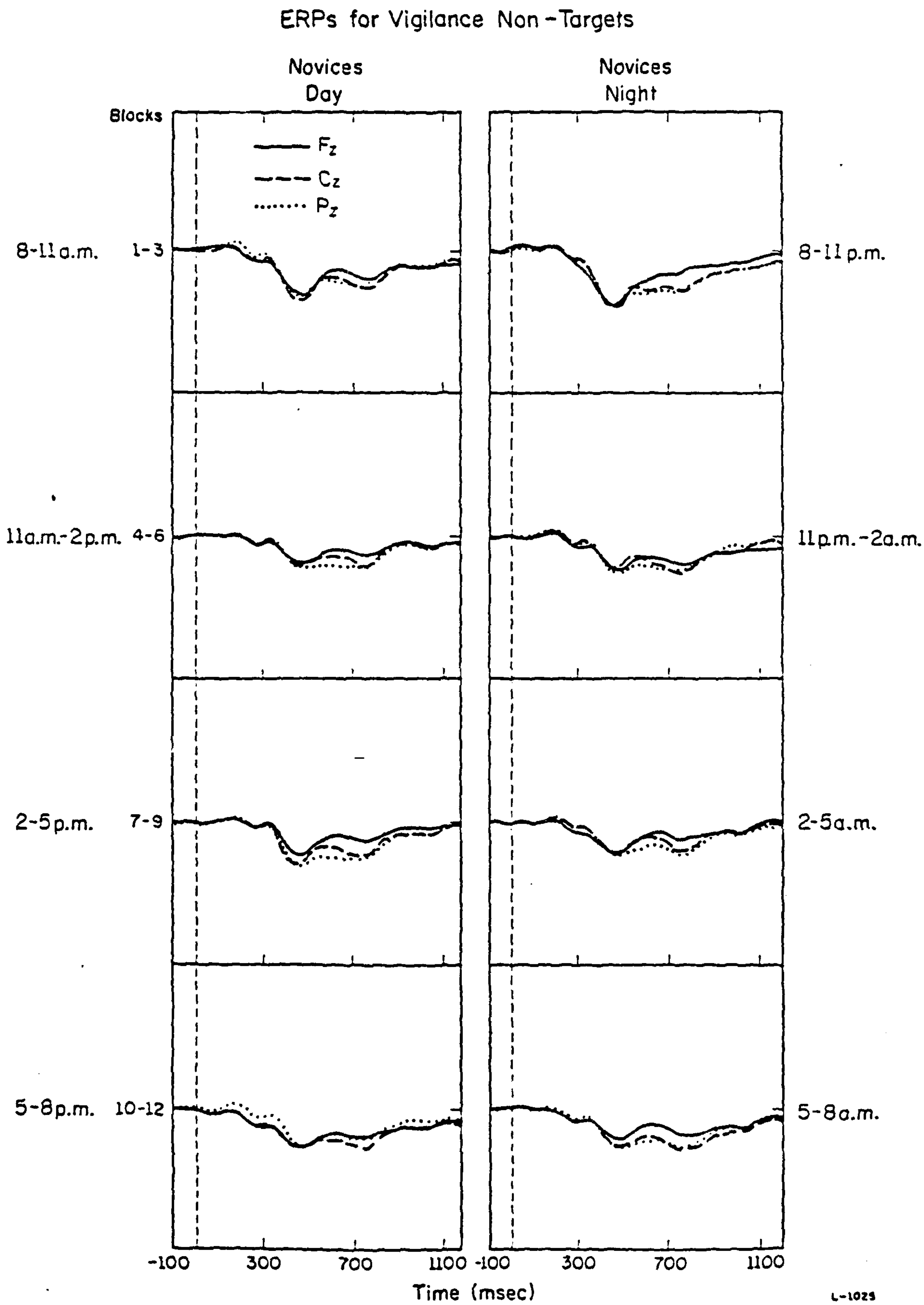
Figure 2. Schematic representation of the sequence of events in the task.

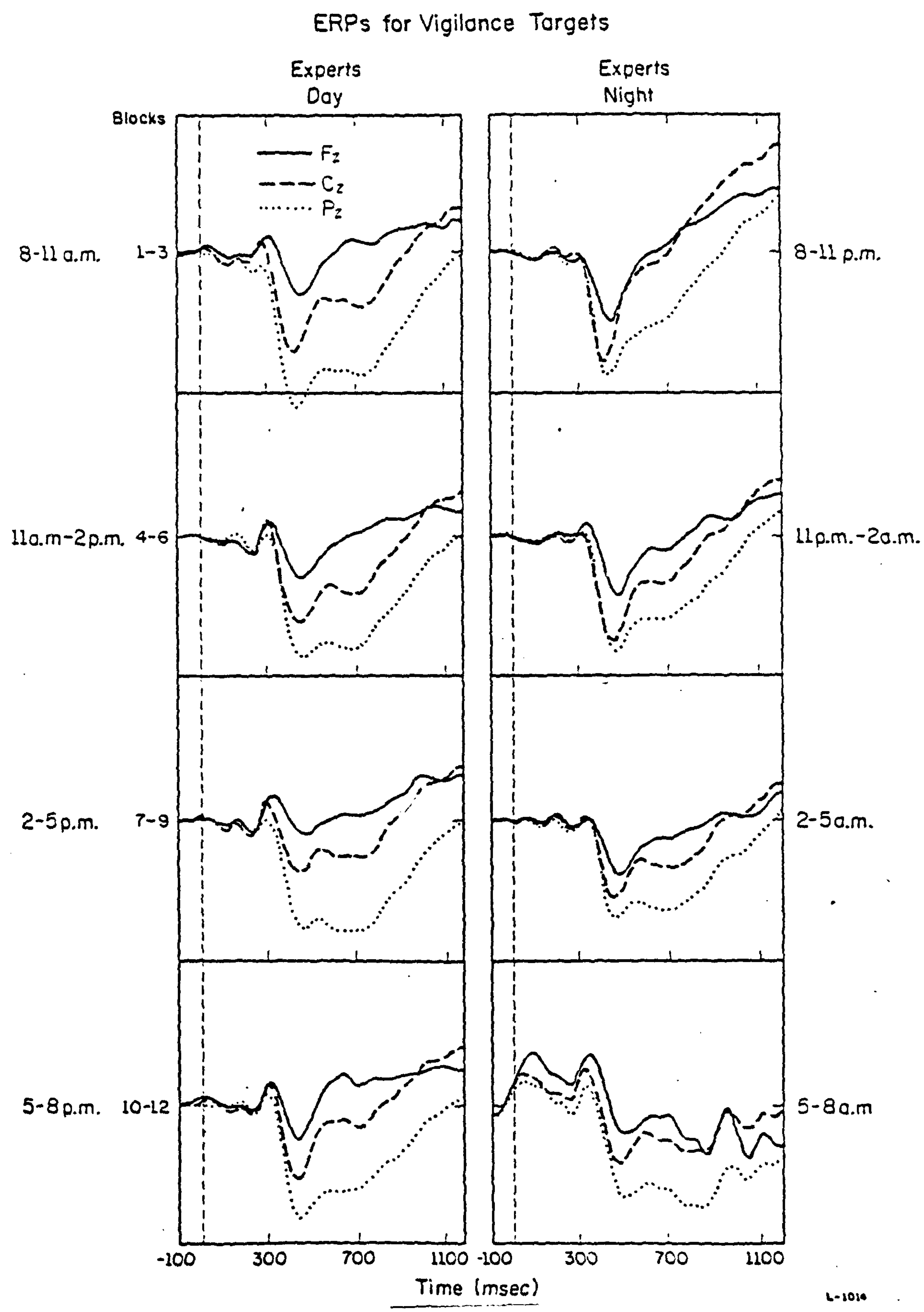
Figure 3. ERPs averaged over 3 consecutive 1 hour blocks. Separate waveforms are presented for experts and novices, day and night shifts, and target versus non-target flashes.


ERPs for Vigilance Non-Targets
Experts Day
Experts Night
Blocks
Fz
Cz
Pz
8-11a.m.
1-3
8-11 p.m.
11a.m-2p.m.
4-6
11p.m.-2a.m.
2-5p.m.
7-9
2-5a.m.
5-8p.m.
10-12
5-8a.m.
-100
300
700
1100
Time (msec)
L-1017

ERPs for Vigilance Targets




ERPs for Vigilance Non-Targets
Novices Day
Novices Night
Blocks
Fz
Cz
Pz
8-11a.m. 1-3
11a.m.-2p.m. 4-6
2-5p.m. 7-9
5-8p.m. 10-12
8-11p.m.
11p.m.-2a.m.
2-5a.m.
5-8a.m.
-100 300 700 1100 -100 300 700 1100
Time (msec)
L-1025



ERPs for Vigilance Targets
Experts Day
Experts Night
Blocks
Fz
Cz
Pz
8-11 a.m. 1-3
11a.m-2p.m. 4-6
2-5p.m. 7-9
5-8p.m. 10-12
8-11 p.m.
11p.m.-2a.m.
2-5a.m.
5-8a.m
-100
300
700
1100
Time (msec)
L-1014

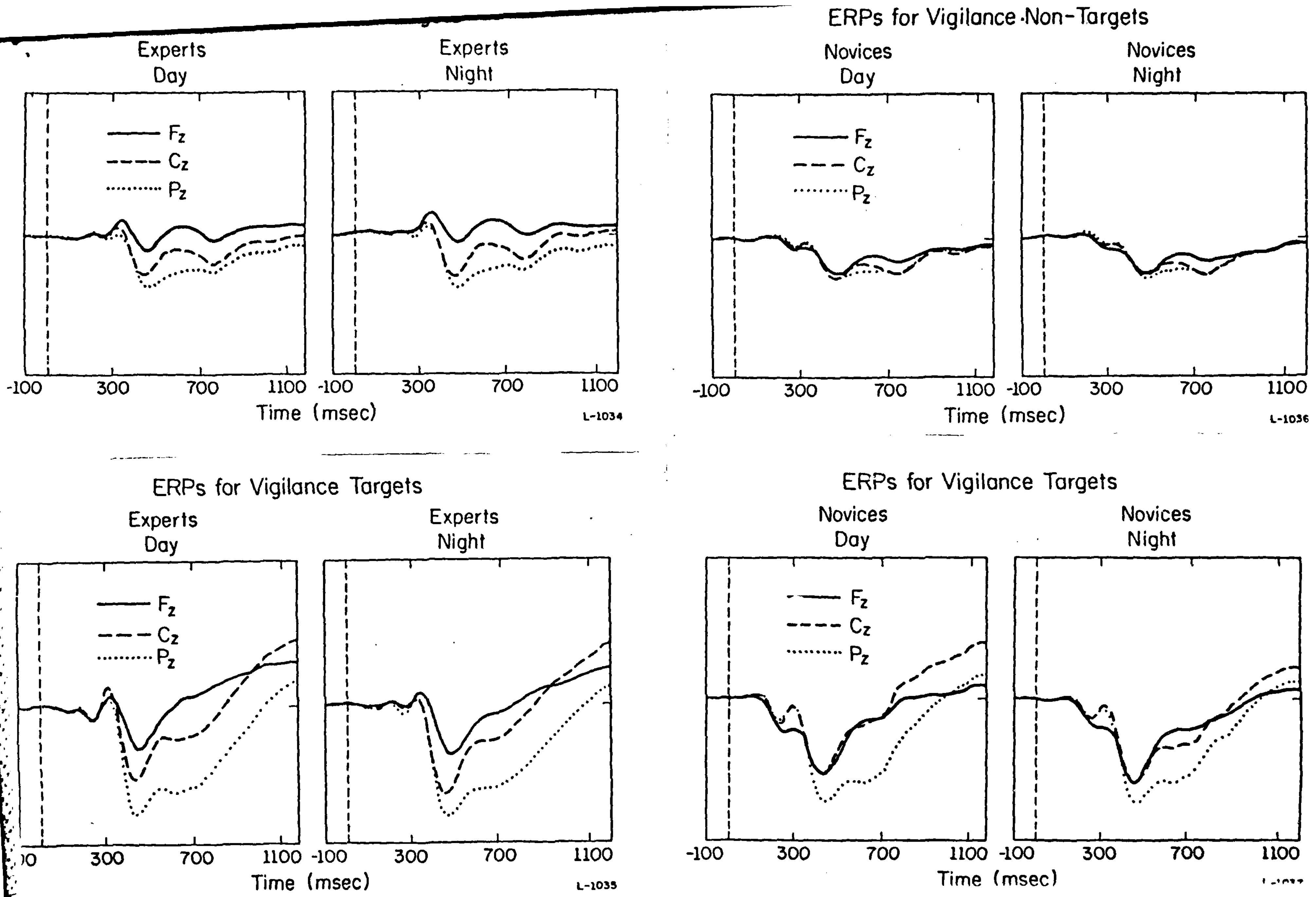


Figure 4. ERPs derived by averaging over the 12 hours of the task.

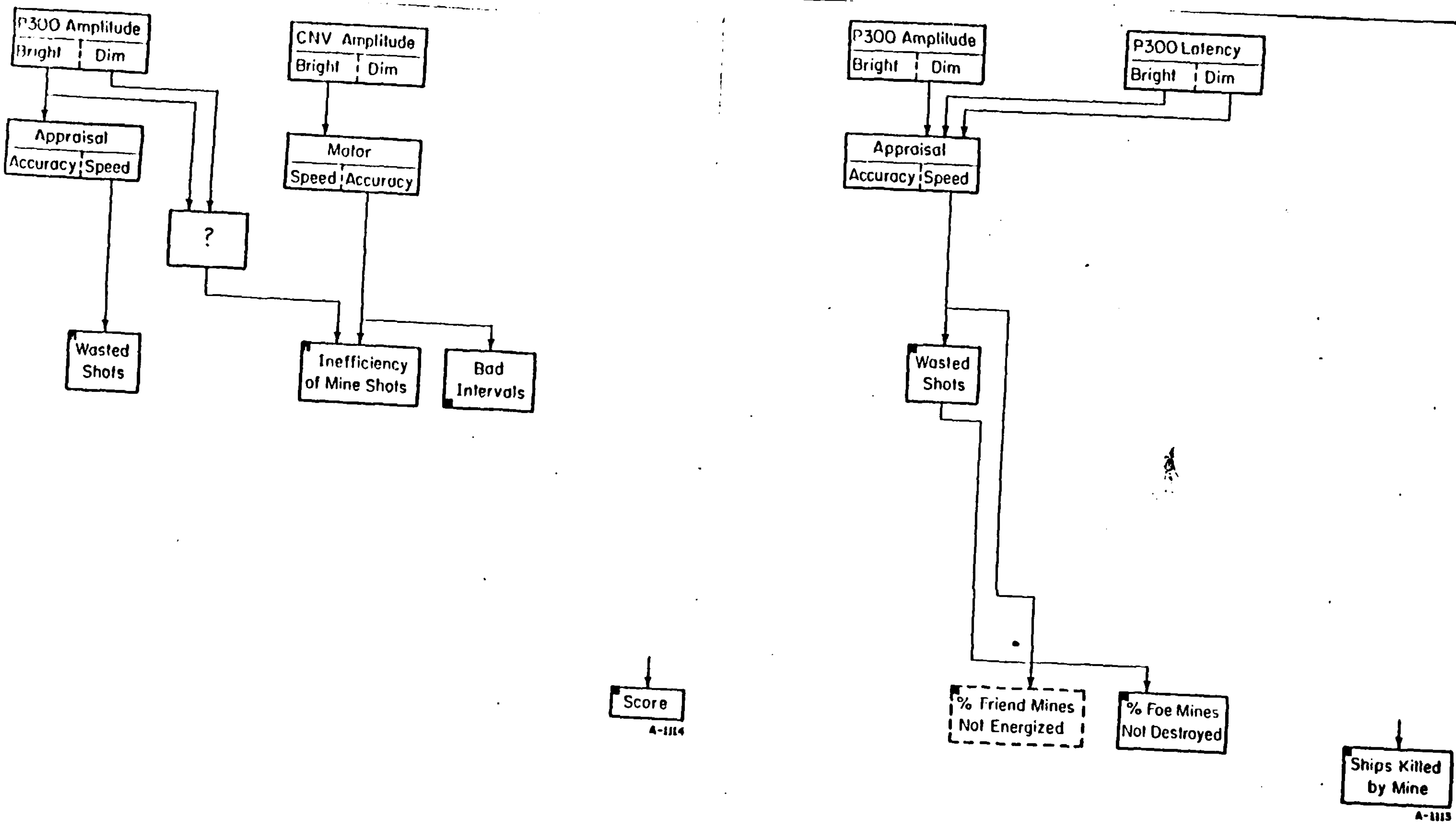


Figure 5. Schematic representation of the relationship between performance and ERP measures - for performance measures derived from (a) Vigilance, and (b) Hell.

APPENDIX B

List of articles supported in whole or in part by AFOSR. Final versions of empirical articles discussed in previous progress reports and review articles and chapters.

1. Polich, J. & E. Donchin P300 and the Word Frequency Effect. Journal of Verbal Learning and Verbal Behavior, in press.

2. Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. Performance of concurrent tasks: A psychological analysis of the reciprocity of information-processing resources. Science, 1983, 221, 1080-1082.

3. Gratton, G., Coles, M. G. H., & Donchin, E. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 1983, 55, 468-484.

4. Donchin, E. The relevance of dissociations and the irrelevance of dissociationism: A reply to Schwartz and Pritchard. Psychophysiology, 1982, 19, 457-463.

5. Karis, D., Chesney, G. L., & Donchin, E. ". . .'twas ten to one; And yet we ventured. . .": P300 and decision making. Psychophysiology, 1983, 20, 260-268.

6. Klein, M., Coles, M. G. H., & Donchin, E. People with perfect pitch process phonic probes without producing a P300. Science, in press.

7. Magliero, A., Bashore, T. R., Coles, M. G. H., & Donchin, E. On the dependence of P300 latency on stimulus evaluation processes. Psychophysiology, in press.

8. Donchin, E. & Bashore, T. R. Clinical versus psychophysiological paradigms in the study of event-related brain potentials. Brain and Behavioral Sciences, in press.

9. Coles, M. G. H., & Strayer, D. L. The psychophysiology of the cardiac cycle time effect. In J. F. Orlebeke, G. Mulder, & L. J. P. van Doornen (Eds.) Cardiovascular psychophysiology: Theory and Methods. New York: Plenum Press, in press.

10. Coles, M. G. H., Gratton, G., Kramer, A. F., & Miller, G. A. Principles of signal acquisition and analysis. In M. G. H. Coles, E. Donchin & S. W. Porges (Eds.) Psychophysiology: Systems, Processes, and Applications. Vol I: Systems. New York: Guilford Press, in press.

END